

Why Do Firms Engage in Open-Source Software Development?

Kevin Mei

October 29, 2025

Job Market Paper

[\[Please click here for the most recent draft\]](#)

Abstract

Open-source development is an unconventional contract where firms can attract non-employees as inexpensive labor. Using task-level data from firms, I examine how the 2022 crypto shock propagates through venture capital relationships to impose financial constraints on other non-crypto startups. Treated firms attract open-source developers as a substitute for hiring, by growing employee headcount 6% slower, posting 15% more public projects, and relying on 10% more contributions from non-employees. The marginal contributors are of lower quality, leading to spillover effects: non-employees participate in open source to benefit from signaling, but signaling-related outcomes decline as firms host more public projects.

McCombs School of Business, University of Texas at Austin. KevinMei@utexas.edu
I thank John Griffin, Mike Sockin, Aaron Pancost, Andrey Ordín, Mindy Xiaolan, Andres Almazan, Aydogan Altı, Jonathan Cohn, Willie Fuchs, Daniel Neuhann, and Sheridan Titman for their generous feedback. I also thank my colleagues Tommy Brown, Minjoo Kim, Eric Lee, Prateek Mahajan, Max Sacher, and Zirui Wang for helpful conversations.

“Microsoft discourages employees from contributing to open-source projects, ... the main effect of this policy may be to ensure that they won’t be able to hire any first-rate programmers.” – Paul Graham, 2003, Y Combinator co-founder¹

1 Introduction

Open-source software is a ubiquitous and invaluable part of the digital world.² In this paradigm, developers release code under licenses that allow anyone to use, modify, and improve it. Open-source is distinctive in comparison to typical economic relationships in that firms share intellectual property with a broad decentralized community without guaranteed benefits, without securing contracts with a specific partner, and without directly compensating contributors. What economic forces sustain these relationships? Under what conditions do firms benefit from this relatively informal, unconventional contracting arrangement? Why would developers volunteer their time? How do firm and worker incentives interact, if at all?

In this paper, I investigate how firms’ project choices interact with workers’ incentives to signal their quality. The core mechanism is that, because high-quality workers benefit from participating in open-source development to publicly demonstrate their ability, firms can offer lower wages for public projects and, in some cases, attract volunteers to contribute. By hosting public projects, firms effectively provide a signaling device to the labor market. However, if firms increase demand for labor in public projects and relatively lower-quality workers begin to participate, a spillover effects leads open-source projects to be a weaker signal of worker quality.

To study this mechanism, I construct a dataset linking three types of information: tasks in firm projects, firm financing history, and worker employment history. For developer tasks, I use data from GitHub, the largest platform for collaborative software development, to study the public projects that have been posted by a firm. In the main sample, I examine

¹From the essay [Hackers and Painters](#): “It seems surprising to me that any employer would be reluctant to let hackers work on open-source projects. When we interviewed programmers, the main thing we cared about was what kind of software they wrote in their spare time...I’ve been told that Microsoft discourages employees from contributing to open-source projects, even in their spare time. But so many of the best hackers work on open-source projects now that the main effect of this policy may be to ensure that they won’t be able to hire any first-rate programmers.”

²U.S. CISA: “Open Source Software [underpins the essential services and functions of modern life](#).” [Hoffmann et al. \(2024\)](#) estimates that if open-source did not exist, then firms would pay \$8.8 trillion to replace their software. [Emery et al. \(2025\)](#) show that the majority of U.S. firms (by market capitalization) release open-source projects and estimates the value of all project announcements at \$25 billion.

the projects and financing histories of over 13,000 firms, from 2018 to the present, and contributions from 1.3 million workers, who submitted 6.8 million tasks, including 32 billion lines of code changed. Within this task-level data, I compare the contributions of “external” developers (i.e., volunteers) and “internal” developers (i.e., employees).³ To evaluate the relevance of signaling, I first test whether developers with more open-source experience earn higher wages in future positions. In a panel regression with firm, month, location, and seniority fixed effects, workers earn 0.9% higher wages in their next role when the share of GitHub tasks completed as external contributors increases from 0% to 50%. Consistent with the signaling channel, I find that the share of a developer’s work made as an external contributor declines with years of experience, or that signaling is more relevant early in a developer’s career. Altogether, these correlations lend credence to the signaling channel, or that high-quality workers exert costly effort with the goal of earning higher wages in the future.

This paper’s primary results use the collapse of FTX, a major cryptocurrency exchange, in mid-2022 as a shock to its venture capital (VC) investors and their portfolios. The sudden, high-profile collapse led multiple VCs to reduce the number of financing transactions, which affects other portfolio companies including firms that are not related to cryptocurrencies. I find that VCs who funded FTX participated in 13% of all deals in 2021, but only 6% of deals in 2023. The treatment group consists of firms other portfolio companies funded by VCs that also invested in FTX. As a test of relevance, a proportional hazard model shows that treated firms were 42% less likely to raise subsequent rounds on a given day. In a difference-in-differences framework, treated firms experienced 6% lower employment growth, which is consistent with being exogenously financially constrained.⁴ How do financially constrained firms manage their projects? I find a 10% increase in open-source work, measured as the share of lines of code changes from external developers. Treated firms also posted 15% more

³Section 3 discusses the delineation between external and internal members. Briefly, each task labels the type of worker who proposed the change to the codebase. Each change must then be approved by an authorized contributor. I use the bottom two levels of authority as external and the remaining three levels as internal. This paper refers to tasks and contributions interchangeably.

⁴Startup executives refer to this practice as managing cash runway, often through reducing burn rate or a RIF (“reduction in force”).

projects publicly, attracted 13% more unique contributors, and produced 17% more code changes to public projects. This is surprising because financially constrained firms typically have worse future prospects, such as in paper like [Campello et al. \(2010\)](#). In heterogeneity tests, I show that the substitution between workers and volunteers is weaker for firms that raised their last round of funding more recently, and that contemporaneous decreases in hiring are associated with more reliance on volunteers. Altogether, these tests demonstrate how financially constrained firms can substitute away from hiring and rely more on the open-source community for contributions.

Do firm reactions have an impact on workers? I measure how the shock impacts the characteristics of external contributors as more firms post public projects. Among the treated group, the marginal external contributors after the shock are of lower observable quality: their cumulative merge rate, or propensity for proposed changes to be accepted into a project, is 1.3% lower, and they are 13% less likely to have contributed to a top-100 GitHub project. This is consistent with the idea that high-quality workers are the “first” to enter open-source projects, and increasing demand for open-source work attracts workers of marginally lower quality. This imposes an externality that open-source work is a weaker signaling device if it does not allow high-quality workers to separate themselves from other workers. I measure this externality by whether a worker is hired, as proxied by whether workers are “promoted” from external to internal. I find that, after the shock, contributors to the treated group are less likely to be promoted both at treated firms and at other firms, even after controlling for observable quality and worker fixed effects.

I explore the channel through additional empirical analysis around worker and firm outcomes. Consistent with the idea that financial constraints are costly, I find that treated firms pay lower future wages and are less likely to IPO. The decrease in IPO likelihood is especially pronounced with firms that had public projects prior to the shock. I find suggestive evidence that firms attract more open-source contributions depending on the project’s primary pro-

programming language. I also examine the career trajectory of employees present at the time of the shock. These workers are more likely to become founders in the future, which may be interpreted as evidence of open-source lowering barriers to entry for entrepreneurship. Further heterogeneity tests analyzes the source of variation that drives the main results, and I find that effects are slightly stronger for small firms. A matching design also produces similar results. I also present additional tests for robustness by varying the assigned treatment group by VC size, a VC’s focus on crypto, and placebo VCs.

The paper then presents an equilibrium model to illustrate how firm project choices interact with worker incentives. The model features a continuum of workers that are heterogeneous in their productivity, and a continuum of firms with two types: financially constrained or unconstrained. The worker types are their private information. There are two distinct markets for public and private work. Firms and workers endogenously sort into either, neither, or both markets. Firms that produce in the market for public work are assumed to have lower output. On the other hand, workers that produce public work receive a benefit that is increasing in worker type, which captures the notion that workers receive higher future wages if their work is public. I focus on a cutoff equilibrium, where the main result is such that workers of higher quality choose to participate in public markets and receive a lower wage, because they can receive higher wages in the future. This model helps rationalize the idea that an increase in demand for public labor will attract more workers of lower quality, which makes the signaling device less effective for hiring.

Overall, the results illustrate how asymmetric information and worker incentives have helped shape the modern digital economy. In this context, because open-source contributions provide a credible signal of talent and human capital is portable, workers have strong incentives to volunteer. Under some conditions, this mechanism supplies firms with free labor. Divergent views on whether open-source software benefits firms have recently resurfaced among leading artificial intelligence startups.⁵ This paper highlights how economic relation-

⁵For example, see “[Open Source AI is the Path Forward](#)” and the discussion therein.

ships through open-source development mitigates asymmetric information in a manner that is mutually beneficial for firms and workers.

Related literature This paper contributes to the literatures on open-source development, relationships between venture capital firms and portfolio companies, and the tradeoff between labor and capital.

My paper engages with a growing empirical literature on how firms benefit from open-source innovation.⁶ [Roberts et al. \(2006\)](#) estimate a simultaneous equations model on open-source software production. [Nagle \(2019\)](#) uses a dynamic panel regression and finds that platforms with complements create additional value-added productivity with open-source. [Petrulia \(2025\)](#) shows how patents and open-source innovation are complements. [Sharma et al. \(2024\)](#) present empirical evidence on how strong complementarities mitigate the substitution effects of open-source in cryptocurrencies. [Hoffmann et al. \(2024\)](#) estimate that firms would need to spend 3.5x more on software if open-source did not exist. Both [Emery et al. \(2025\)](#) and [Coleman et al. \(2025\)](#) find that releasing open-source software is associated with positive stock returns and higher future earnings. While [Emery et al. \(2025\)](#) also finds that more permissive licenses are less valuable and product complement is not a primary driver of value, in contrast, [Coleman et al. \(2025\)](#) highlight that open innovation complementarity is a determinant of establishing a presence on GitHub and is also associated with higher risk of cybersecurity breaches. While these papers focus on the firm’s decision to engage in open-source software development, my paper emphasizes the role of labor in open-source development: successful open-source projects require a strategic interaction between the firm’s decision and labor’s decision to participate in open-source. Other papers that emphasize labor include [Hakim Orman \(2008\)](#), who uses an Arora-Gambardella model to show that open-source is a complement to education. [Gortmaker \(2025\)](#) estimates a structural model

⁶Researchers also provide a perspective on why developers prefer open-source software. [Roberts et al. \(2006\)](#) studies the Apache projects, a collection of frameworks that power most of the Internet experience, and find that reputation and past-performance rank high among engineers’ motives. [Lakhani and von Hippel \(2003\)](#) survey software engineers and study a message board to find that reputation is strong self-reported motivation, as well as altruism. Pay ranks low. Many prominent essays from the software industry have also argue that open-source projects have higher quality. Linus’s Law posits “given enough eyeballs, all bugs are shallow.” Paul Graham echoes “open source software is more reliable precisely because it’s open source; anyone can find mistakes.”

of open-source production and labor with spillovers. [Boysel et al. \(2024\)](#) study whether labor market competition has an inverted-U relationship with the number of open innovation contributions, in the spirit of [Aghion et al. \(2005\)](#). [Abou El-Komboz and Goldbeck \(2024\)](#) define “active” job seekers and observe that they make more open-source contributions. My paper studies open-source through the lens of how the firm’s decision interacts with worker incentives. I bring causal evidence to study how firms can use open-source to alleviate financial constraints by relying on the worker’s incentives for free work.

Empirically, I build on an identification strategy that considers how investor constraints impact their portfolio firms. [Townsend \(2015\)](#) uses the 2000 dot-com bubble and finds that other firms that were exposed through their investors were delayed in receiving funding.⁷ I build on this methodology and apply it to the 2022 crypto collapse to study how a shock to financing conditions impacts a firm’s projects and labor decisions.

This paper studies open-source from the lens of signaling, as espoused by [Lerner and Tirole \(2002, 2005\)](#). The hypotheses I test are similar in spirit to theoretical results in [Ke et al. \(2023\)](#), which models a career concerns à la [Holmström \(1999\)](#) with the choice of an informative or a productive task and predicts that a worker chooses more informative tasks when their reputation is low. [Kumar et al. \(2011\)](#) and [Leppämäki and Mustonen \(2009\)](#) explicitly model signaling in open-source and their spillovers. As a contribution, this paper explicitly models the coexistence of private work by employees, open-source work by employees, and open-source work by volunteers.⁸

My results speak to the relationships between capital and labor. [Michelacci and Quadriini \(2009\)](#) model a labor market where financially constrained firms borrow from their workers by promising higher wages in the future. [Eisfeldt and Rampini \(2007, 2008\)](#) study new-or-used-

⁷Other papers studied how shocks to VCs propagate to their portfolio companies. [Howell et al. \(2020\)](#) studies how innovation at VC-backed startups pro-cyclical. [Nanda and Rhodes-Kropf \(2013\)](#) show how timing of VC investments is associated with riskier investments and more successful patents. [Zandberg et al. \(2024\)](#) find that investments made early in a VC fund’s lifecycle is associated with stronger monitoring. [Eldar and Grennan \(2023\)](#) considers how lifting restrictions on VC investments impacts portfolio company competition. [Ersahin et al. \(2021\)](#) show how increasing VC competition can crowd out high-quality VCs. [Chen and Ewens \(2025\)](#) shows how VC financial constraints spillover to startup geographic agglomeration. [Bernstein et al. \(2022\)](#) show that the presence of high-quality VCs can attract more talent. [Gu et al. \(2017\)](#) show how greater labor market frictions in a state lead to lower VC investment and lower rate of patents.

⁸Papers like [Llanes and de Elejalde \(2013\)](#) present theories where firms benefit from open-source software through complementarities.

capital and buy-vs-lease decisions, respectively, and highlight how financially constrained firms defer cash outflows. [Garmaise \(2007\)](#) develops a model where financially constrained firms use more labor than physical capital. My paper studies a similar trade off in the firm’s decisions. [Sun and Xiaolan \(2019\)](#) highlight how firms use employee financing through delayed wages in the form of equity claims. I contribute by studying this from the angle of open-source, a contracting framework where firms do not directly pay nor contract with workers, and by bringing new empirical evidence on project choices.

I first develop hypotheses in [Section 2](#) to guide my analysis. [Section 3](#) describes the dataset linking firms and developer histories and presents summary statistics and panel regressions to explain the data. [Section 4](#) presents empirical results from an exogenous shock in a difference-in-differences framework. [Section 5](#) includes robustness tests. Lastly, [Section 6](#) presents an equilibrium model to formalize these findings. [Section 7](#) concludes and includes a discussion on why mitigating asymmetric information is particularly relevant in software development, a critical industry in the modern economy.

2 Hypothesis Development

This section develops hypotheses on the incentives for open-source development. I describe tradeoffs first for firms, then for workers. Consider a firm that has an early prototype of a project and requires additional workers. A firm has at least three options: (1) hire labor to complete the project privately; (2) hire labor to complete the project and share project’s code with the broader public; or, (3) first share the prototype with the broader public and find a developer to volunteer their time and complete the project. The act of publicly sharing code, or any intellectual property, may increase or decrease a project’s profitability.⁹ This paper focuses on the firm’s tradeoff where sharing code online is costly because it reduces profits from the project. The benefit to open source, however, is that firms can lower labor expenses, especially if volunteers complete the project as in Option 3.

⁹One potential reason an public project reduces profits is because it potentially allows competitors to copy the firm’s intellectual property that would otherwise be exclusive. Another reason may be if the project is delayed when waiting to find a suitable volunteer, or if the firm cannot sufficiently monitor a volunteer. Alternatively, a firm may also benefit from sharing code online, such as if sharing code increases profits by driving business to complementary lines of business. This paper does not focus on these cases because the solution would be such that firms always prefer sharing code online.

Worker incentives also interact with firm choices. Many labor markets feature workers with vast heterogeneity in ability and employers who are unable to easily find the best workers within the pool. A classic resolution is [Spence \(1973\)](#) signaling where workers exert effort to produce a costly signal that separates high-ability workers from the rest of the pool. Open-source development can be a useful signaling device for high-quality workers to clearly demonstrate their ability because workers can post code online.

When weighing these three options together, workers should prefer to work in firms that host public projects because it allows an individual worker’s output to be visible online, and improves the value of the worker’s human capital. This produces a natural hypothesis:

Hypothesis 1. *Firms trade off between hiring costly labor for private projects or seeking volunteers on a public project. Because workers place higher value on public projects, firms may be able to pay lower wages on a public project, and may allocate tasks towards public projects when financially constrained.*

Firms with public projects effectively produce signaling devices for workers. While workers can also produce their own public projects, it may be that working on a firm’s projects is a stronger signal. This is related to the core Spence mechanism: signaling arises from higher-type agents separating themselves from other types by exerting costly effort, and the specific shape or quality of effort does not matter so long as it is too costly for a lower-type agent to imitate. Therefore, the provisioning of signaling devices also carries an externality:

Hypothesis 2. *If firms increase demand for labor in public projects, and under conditions where the average quality of workers in public projects declines, then signaling is less informative because high-quality workers cannot meaningfully separate from other workers.*

In [Section 6](#), I develop a model where firms and workers sort into public and private projects. Because future employers want to hire high-quality workers, the high-type can signal their quality by taking an action that would be too costly for a low-type worker. In this setting, high-type workers forgo wages by working on public projects. However, within

this mechanism, I highlight how signaling becomes less useful if it is too common.

3 Data and Summary Statistics

This section describes the data and approach to merging data on firm projects, financing, and labor. I first discuss the rationale behind how the sample was constructed, then describe institutional details about open-source development. I lastly provide summary statistics and basic statistical relationships in the data.

I construct a dataset by linking data from GitHub, Pitchbook, and LinkedIn in three main steps. First, from GitHub, I gather organization names, websites, and locations listed for projects with more than 50 stars.¹⁰ I aggregate GitHub owners based on their organization ID. The benefit of this approach is that IDs are immutable even as organizations change their name. This sample also mitigates survivorship bias by collecting all websites over time including any websites that are now defunct. I exclude organizations that list their website as GitHub or other social media accounts and remove duplicates, which yields 55,646 firms. I rely on PitchBook as my primary sample of company information and fuzzy match the GitHub sample using company name, website, and location.¹¹ Lastly, PitchBook data includes LinkedIn profiles, which I use to match against data from Revelio Labs, yielding 14,667 matches. This sample construction approach centers on PitchBook’s internal database, which has the advantage of being a well-defined set of firms with cleaned data on websites, LinkedIn profiles, locations, and industries.

I use data on GitHub activity of developers from GH Archive.¹² This dataset collects all user actions from the GitHub Timeline API, including stars, forks (copying a project), pull requests (proposing edits to a project), and other user actions. In GitHub, developers are granted permissions to a project in five tiers. I count the top three tiers to be “internal” and the bottom two as “external” developers. These designations are meaningful because they

¹⁰I limit the sample to focus on projects with substantive interest. There are over 403,202 projects with more than 50 stars, which I fuzzy match within PitchBook. Users can “star” a project or firm to save it as a bookmark and follow its progress.

¹¹I only retain 28,257 matches with *Very High* or *High* confidence (23,996 are matched with *Very High*) from PitchBook’s matching algorithm. I drop searches that yield *Medium* confidence or no match.

¹²Interested readers can learn more about this dataset: <https://github.com/igrigorik/gharchive.org>

grant the developer access to specific actions that are important for the functioning of the project. For example, proposed code changes must be approved by a project member that has proper authority. On the GitHub platform, a project’s codebase hosted by an organization is referred to as a *repository*. Proposed code changes to a repository, known as a *pull request*, are the most granular task or unit of work in the scope of this paper. Within pull requests, I measure total work as the sum of the number of *additions* and *deletions* of lines of code. Developers commonly call these *diffs* and, throughout the paper, I refer to this as the volume of code changes. *External merged share* is then the share of merged pull request additions and deletions completed by an external developer. To focus on worker actions, I exclude “bots” from the sample based on their GitHub label and additionally exclude account names that use “_bot” as a suffix.

Figure 1 presents a cross-sectional summary as of the end of 2024 on the volume of accepted code changes and share of work by external developers. Several striking patterns emerge. Organizations often consist of 0% external contributions or 100% external contributions, as is visible by the darker density near the top and bottom of the figure. However, a sizable number organizations are also within the middle of the distribution, especially larger firms towards the right along the horizontal axis. Many of the world’s largest corporations also appear on this page. These dots represent only the public projects that have been published to GitHub. There is large variation for firms within the same industry. For example, Intel appears near the top with 80% external, while NVIDIA appears with less than 20%, and AMD is somewhere in between. Moreover, volunteer share has only a -0.01 correlation with merged code changes, and -0.02 with the number of followers, suggesting that size and popularity alone do not drive open-source engagement. Among the large companies, it is clear that more stars (as indicated by the size of the dot), is not highly correlated with more external contributions. Lastly, as institutional background, it is worth noting that these projects are also very heterogeneous in business objectives. Some projects are strong complements to the core lines of business of each firm, while others appear unrelated to the

firm.¹³

Figure 2 presents summary statistics on the volume of work completed by contributors in the five levels of author associations. These are assigned as permissions by the GitHub owner. A contributor is anyone who has a history of completing a task (i.e., their pull request was merged into the codebase), while collaborators are those that are explicitly granted “write” access to a project. On the left-hand side, we see that most users have no status. Collaborators have the largest median and 90th percentile of contributor size. Contributions that are merged, or accepted, tend to be smaller than rejected contributions.

Table IA.I presents the total number of pull requests and lines of code changed. The acceptance rate of tasks declines as developers are of lower permission levels. The largest volume amount of work is completed by collaborators and contributors. However, collaborators propose code changes with larger volume and are more likely to be accepted.

To better understand the differences between internal and external developers, I use machine learning to compare code contributions. I use batch chat completions with the large language model (LLM) 4o-mini by OpenAI to rate the level of sophistication on a scale of 0-10. Figure IA.1 plots the distribution of ratings with internal developers in blue and external developers in orange. Although the t-test suggests that external developers are marginally less sophisticated, the distributions are similar. As a whole, these summary statistics suggest that code length and code quality are similar, and that the intensive margin on share of work by external developers is unlikely to be rationalized by a trivial explanation such as if external developers only make cosmetic code changes.

My main sample contains approximately 13,000 firms where I can merge financing from PitchBook, firm-level public development activity from GitHub, and employee resumes from Revelio Labs. To link GitHub organizations to their listed employees on LinkedIn, I use PitchBook data on each firm’s LinkedIn profile and merge this with the GitHub-PitchBook

¹³For example, while Palantir is known for data analysis solutions, their most active project is related to front-end development. Other firms are more closely related: Facebook has released important libraries for front-end development and machine learning. Netflix is well-known to have a strong culture of developer operations (“dev ops”) and has released many libraries to test system reliability.

sample. Further, when considering the employment of software developers, I focus on Revelio profiles that have any history of roles that are tagged as 15-1252.00, the O*NET code for Software Developers. This yields 3,216,495 workers for whom I collect their employment history to track their career progression.

I use other data on users and their positions from Revelio Labs as accessed through WRDS. I use the data field of worker compensation, which is predicted by Revelio and is not definitive compensation data. Based on online documentation, Revelio predicts compensation using a mix of worker and company reported data and accounts for covariates such as location and seniority.¹⁴ I use this field primarily to compare wages across workers and include fixed effects for location and seniority whenever applicable. I also use education and label top schools as the top 100 universities, based on their graduates’ prestige, which is also a data field ranked by Revelio. Data on job locations is used to study workforce composition. I also extract data on whether a worker is a founder by checking titles for instances of “founder,” which nests “co-founder,” or “founding engineer.” Similarly, I extract data from job descriptions to flag if the text mentions “open-source” and its variants.

3.1 Empirical Facts

I first document empirical facts related to the signaling mechanism in open-source development. The goal of this subsection is to determine whether basic predictions from signaling models are relevant in the data. For the analysis in this section, I use GitHub-LinkedIn matches to link 16,865 developers to their employment. I make a strong assumption that developers use the same login name for GitHub and LinkedIn and manually verify this assumption for a subsample. This is reasonable because the default generated LinkedIn login is a combination of first name, last name, and digits, which prompts many users to change this to their preferred online handle.

One interpretation of signaling models predicts that workers will invest in signals early in their career and reap the benefits later. Figure 3a explores the relationship between

¹⁴More information on Revelio methodology [is available here](#).

years of experience and share of a developer’s work that appears to be volunteer effort. The binscatter includes developer fixed effects to absorb idiosyncratic preferences for open-source work. With less than 5 years of experience, more than 50% of a developer’s contributions are made as an external contribution; however, most of the data are from software engineers in the 5-15 years of experience range, where the average is between 40-45% of work. Consistent with a signaling channel, volunteer effort appears to decline as workers accumulate more years of experience.

Do developers who make unpaid contributions earn higher wages in the future? I build a panel dataset where the unit of observation is a worker-position and present results from a fixed effects regression in Table 2. The first column includes fixed effects such that the specification compares the future wages of workers after removing variation from the firm, start month, location (defined as state if in the U.S. or country otherwise), and seniority (as defined by Revelio). Comparing a worker whose contributions in GitHub were 0% to one with 100% as an external member, the worker with more volunteer effort earns 1.9% more in the next job. Column 2 includes developer fixed effects such that absorb the idiosyncratic ability and preferences for a worker. The coefficient is similar in this comparison of different positions within the same developer’s work history. In summary, when a worker continues to make contributions to projects where they are not a collaborator, then they earn higher future wages.

I also test for the presence of strictly increasing differences in types, or that it is relatively cheaper for high-type workers to generate a signal, a common assumption of Spence signaling models. If so, then high-type workers can also signal for “longer” in their career than low-type workers. Empirically, high-type developers should make more open-source contributions than low-type developers as years of experience increase. I test this by using education as a proxy for high-ability workers, and formally check if workers from top universities make more open-source contributions, and for longer in their career. I designate *top schools* as those

ranked in the top 100 universities globally. Graduating from a top university is already a signal of high ability for these workers. Nevertheless, workers may still seek additional signaling devices to further separate from other top graduates. Signals from open-source development may also capture other skills that are more relevant than education.

Table [IA.II](#) presents results that, in the cross-section, developers from top schools tend to make 3% more external contributions but only in the first five years of their career (Column 1). This effect disappears in the full sample (Column 2). Similarly, as years of experience increase, external contributions decline (as predicted). In Column 4, developer and date fixed effects absorb the variation from the years of experience variable, but a positive coefficient for the interaction with top schools suggests that workers from top schools contribute more open-source compared to all other developers as experience increases.

The prior regressions establish multiple statistical relationships that are consistent with the signaling channel for open-source development. Using a subsample constructed from matching GitHub and LinkedIn profiles, I show that workers exert costly effort through open-source development and are predicted by Revelio to have higher wages in the future. The variation in the data mostly comes from early career workers, and “high quality” workers, as proxied by having attended a top university, although there are also strong individual fixed effects. Altogether, these results establish that workers have an incentive to participate in open-source development as a signaling device.

4 Main Results: Firms Substitute Workers and Volunteers

In this section, I investigate my hypotheses in a causal setting. I first briefly describe the endogeneity problem. This section then describes the institutional background for the causal shock, including a discussion on timing and establishing that the shock is relevant for firms, before presenting the paper’s main results.

While the previous section documented that workers benefit from engaging in open-source activity, the panel regressions show only suggestive evidence of a tradeoff where firms com-

compensate workers with signals instead of wages. The ideal experiment would be if some firms exogenously increase their proclivity for open-source development, and then an econometrician measures whether treated firms attract workers that accept a wage at a discount to their true quality because workers have a preference for open-source. Lacking the ideal experiment, a panel regression of open-source and wages suffers from endogeneity concerns. The main threat to identification is omitted variable bias because the econometrician cannot perfectly observe firm and worker characteristics. For instance, a worker may accept lower wages not because of the benefits associated with joining an open-source project, but because they expect the firm has strong future prospects. The econometrician also does not know true worker quality and thus cannot assess if workers accept a discount wage in exchange for open-source projects or any other firm characteristic. An exogenous shock to incentives is needed to disentangle these effects. This section presents an exogenous shock to the financial constraints of startup firms, which causes firms to reallocate their project labor according to their incentives.

Treatment and Timing

I use the unwinding of FTX as a shock that propagates to startups through overlapping financiers. Between April and June 2022, Bitcoin prices fell by approximately 60% and Ethereum declined 70%. The drawdown in cryptocurrency valuations weakened FTX, which later led to its collapse. The treated group are startups that received financing from the same VCs as FTX as of Q2 2022. A total of 83 VCs had invested in FTX as of the time of their collapse. Financiers included a diverse range of firms ranging from large high-profile firms, such as Sequoia Capital, to individual investors. Many financiers of FTX publicly exited the crypto industry, including Tiger Global, Singapore sovereign wealth fund Temasek, Ontario Teachers' Pension Plan, and SoftBank Vision Fund.¹⁵ As the problems at FTX became apparent, industry participants and fund beneficiaries began to cast doubt on the due diligence processes of FTX's institutional backers.¹⁶ This is a source of exogenous variation

¹⁵FT: Ontario Teachers fund steers clear of crypto after \$95mn FTX loss; Forbes: SoftBank Puts Blockchain Investments On Ice As Part Of Startup Pullback;FT: Singapore's Temasek cuts back on start-up investments

¹⁶NYT: Investors Who Put \$2 Billion Into FTX Face Scrutiny, Too

to the treated group in that the fallout from the collapse was unpredictable, especially because the magnitude of the problems in FTX was unknown to the public prior to this time. Robustness tests will vary the treatment group and show that results are not driven by outlier subsets of VCs. Overall, the treated group comprises 12% of my sample and 12% of the sample is related to cryptocurrency.¹⁷ The interpretation is that, when facing a tight funding environment, these portfolio companies would expect that the next round of funding is delayed and management teams will conserve cash.

The timing of this shock is in a similar spirit to papers such as [Ivashina and Scharfstein \(2010\)](#) and [Chodorow-Reich \(2013\)](#), which use bank exposure to Lehman Brothers, in that the financing constraints tightened over time. In the Lehman case, as [Ivashina and Scharfstein \(2010\)](#) write, “the main problems in this market were already evident almost a year before Lehman’s failure,” which included Bear Stearns failing in March 2008 before Lehman failed in September 2008.¹⁸ The crypto collapse similarly unfolded gradually. Cryptocurrency returns featured high betas starting in 2020 and so when the broader market faced a -25% drawdown in 2022, the blockchain industry suffered significantly. In June 2022, Coinbase announced an 18% reduction in headcount as the CEO said “we appear to be entering a recession... could lead to another crypto winter.”¹⁹ As other startups collapsed, FTX grew entangled in court proceedings to resolve creditor claims, and the magnitude of FTX’s problems grew apparent.²⁰ By September 2022, there were public clues that FTX’s proprietary trading desk, Alameda Research, generated abnormal profits from the exchange. FTX faced a run and declared bankruptcy in early November 2022.

I find that financing started to dry up after the second quarter of 2022 and therefore use dates after Q2 2022 for the *post* indicator variable. Figure [5a](#) shows new deals began to decline in June 2022 and FTX’s investors’ participation declined. In 2021, prior to the

¹⁷I use PitchBook’s list of cryptocurrency-related firms to first find their current focus group for the sector. I then filter the blockchain-related keywords assigned to these firms to find other firms with the same keywords for my final list of firms in the cryptocurrency industry.

¹⁸[Townsend \(2015\)](#) also discusses the exact timing of the shock from the pop of the dot-com bubble.

¹⁹[WSJ: Coinbase to Lay Off 18% of Staff Amid Crypto Meltdown](#)

²⁰[Bloomberg: Crypto Quant Shop With Ties to FTX Powers Bankman-Fried’s Empire](#). “Between June 1st and July 22nd, Alameda’s known wallets were the largest stablecoin depositors and sources of liquidity to all of FTX’s known wallet addresses”

FTX collapse, these investors were in approximately 13% of deals. After the collapse, they are featured in just 6% of deals in 2023. To formally test a delay in funding, I build on the approach in papers like [Townsend \(2015\)](#) and estimate a Cox proportional hazard model of the form:

$$h_{is}(t) = h_0(t) \exp\left(\beta_1 1(\text{Post})_{is} + \beta_2 1(\text{Treat})_i + \beta_3 \left(1(\text{Post})_{is} \times 1(\text{Treat})_i\right) + \sum_r \gamma_r \mathbf{1}\{\text{VC Round} = r\} + \sum_j \delta_j \mathbf{1}\{\text{Industry} = j\}\right) \quad (1)$$

The results are presented in Table 3. The main coefficient of interest is the interaction term, which ranges from -0.298 to -0.551, depending on inclusion of round and industry fixed effects (Column 2), triple-interactions with whether the firm is in the crypto industry (3) and a subsample of just crypto firms (4). The interpretation of the raw coefficient of -0.541 is that a firm is 42% less likely to receive funding on any given day.

Firm and Worker Responses

Having shown that related firms are plausibly financially constrained, I next turn to examine how treated firms respond. Hypothesis 1 considers how firms can use open-source development to save costs. In this setting, for example, if VCs communicate to portfolio firms that the next funding round will be delayed, then the startup managers may shift to rely more on open-source development as an opportunity to manage expenses. I test if firms substitute away from paid workers and towards open-source development in the following difference-in-differences framework and present results in Table 4.

$$y_{i,t} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{Treat})_i) + \chi_i + \delta_t + \varepsilon_{i,t}$$

Columns 1-2 show employee growth rate declines 6.2 percentage points, or 2.4 percentage points after including industry and geography fixed effects with post-interactions. A decline in headcount growth is consistent with the interpretation that constrained firms will manage

labor expenses to reduce costs. In Columns 3-4, after the shock, the share of accepted work completed by external developers increases 10 percentage points.²¹ Firm fixed effects absorb variation innate to a firm, such as if firms have an innate preference for open-source or a “hacker” ethos. The result is robust to controlling for whether the firm is in the crypto sector, suggesting that this shock is not specific to crypto sector. While columns 1-2 showed a large decline in the coefficient when including controls, which suggests that a large portion of the effect can be explained by observable characteristics, columns 3-4 showed a small change when including controls and a noticeable increase in R^2 . This suggests that the shock is less correlated with observable characteristics when compared to columns 1-2. I also plot the dynamic coefficients on Figure 6 to check for pre-trends. The effect appears starting in the third quarter of 2022, which is when the same VCs as those investing in FTX appeared to participate in fewer deals.

To further investigate the substitution between workers and volunteers, I briefly examine two important triple difference-in-differences tests. First, I use an indicator of whether a firm has more than 500 days since the prior round of funding, which is a round number near the mean 485 number of days between funding rounds. This variable introduces more exogenous variation because the time of last funding is plausibly independent of when FTX collapsed. The null hypothesis is that this has no impact on employment growth or share of work by volunteers, while the alternative hypothesis is that firms that were last funded more than 500 days ago would need more funding and are likely to substitute volunteers for workers. As seen in Table IA.IV, treated firms that last received funding more than 500 days had even lower headcount growth after the shock, and substitute more to volunteers. Next, I also examine whether firms that experience more decline in headcount also shifted to increase share of volunteers. A second triple difference-in-differences test interacts the variable of interest with contemporaneous headcount growth. Although interacting the difference-in-differences test

²¹This regression uses a panel dataset where the unit of observation is a firm-quarter. It is weighted by volume of code contributions so that it represents the average line of code written across all firms. I use weights because subsequent analysis also consider a panel regression at the more granular project-quarter level, and weighting allows the estimate to be agnostic to the level of aggregation. As a robustness check, Appendix Table IA.VI also presents an alternative specification where the dependent variable is log volume of code changes by external developers and shows that total volume also differentially increases for treated firms after the shock.

with an outcome variable is not exogenous, this tests is still useful in examining if firms in the sample that reduce headcount are the same firms that increase volunteer share. Table [IA.V](#) shows a negative coefficient, which validates that firms substitute between headcount growth and volunteer share.

I examine other empirical evidence on whether firms shift towards more public projects by applying a similar difference-in-differences specification and change the outcome variable to log cumulative number of projects, log cumulative number of unique contributors, and log cumulative volume of code changes. I find that number of public projects increases by 13% and volume of code changes increases by 17%, as presented in Table [5](#). The overarching result is that total public projects increased. I also find that treated firms gain 13% more unique contributors and “stars” grow by 17%, or that the projects become more popular with users. This is also surprising because financially constrained firms typically have poor future prospects. Comparing these results with the prior result on headcount decline yields new implications. A decline in headcount growth is consistent with the idea that labor supply will respond to the financial conditions of the firm, as in [Titman \(1984\)](#) and the literature summarized by [Matsa \(2018\)](#). However, while paid workers may leave a distressed firm, my results add a new angle to this literature in showing that unpaid labor does not face the same financial concerns, and treated firms still attract labor through open-source projects.

My results show labor growth and external share appear with opposite signs after facing a shock to firm incentives. Importantly, this is not the relationship present in a panel regression: external share is positively correlated with headcount growth, as shown in Appendix Table [IA.III](#), even after accounting for firm and date fixed effects.^{[22](#)} One interpretation of the panel regression is that healthy firms engage in open-source development, and they can hire employees who also prefer a company with open-source projects. However, my results show a negative relationship because I use a causal setting after a shock to firm financial

²²Appendix Table [IA.III](#) shows that hiring rate is positively correlated with external share, while separating rates have no relationship. The net effect is that headcount growth is associated with external share.

situations. I interpret this as evidence that this shock is meaningfully demonstrating the tradeoff from Hypothesis 1, and that firms substitute between compensating workers with wages and allowing workers to use public projects as signaling devices.

Hypothesis 2 suggests that a firm’s decisions to pursue public projects have externalities for workers. I test whether firms’ pursuit of additional public projects leads the marginal external developer to be of lower quality, and if this reduces the signal associated with open-source development. First, I use two proxies to measure quality: a developer’s merge rate, or the share of contributions accepted, and a binary variable on whether the developer has had a contribution accepted to a top 100 organization, ranked by star popularity. I use the same difference-in-differences framework to test whether the characteristics of the volunteer workers change after the shock for treated and control firms and present results in Table 6. External contributors after the FTX shock have a 1.3 percentage point lower merge rate and are -13% less likely to have made a top contribution. This suggests that the marginal volunteer that firms attract is of lower quality. I rationalize this finding in a model in Section 6.

If the marginal worker is of lower quality, then the signal associated with open-source development may be less effective. I define effectiveness as whether the signal improves a worker’s career progression, such as by being hired. To measure career progression, the indicator variable $1(Promoted)$ is encoded as 1 if, within the next six months, a worker previously had GitHub permissions of “no status” or contributor and later rises to collaborator, member, or owner for an organization. This is a proxy for a worker being hired. I build a panel dataset of all contributions by developers who contributed to any GitHub firm successfully matched to PitchBook. The variables of interest are the amount of monthly work contributed to treated and control firms as external developers, and the outcomes of interest are whether workers were promoted to internal developers at this organization or any other organization.²³

²³I use this dataset to gather a panel of all contributors to treated and control firms. I do not use the dataset that links individual

Table 8 first presents results on being promoted at any organization in the next six months. All columns include worker and date fixed effects, as well as merge rate and total contributions, both interacted with a 1(post) indicator, to additionally control for worker observable quality. In the first row, a 10% increase in contributions as an external developer to a treated firm increases the rate of promotion by 0.07 percentage points (based on the coefficient of 0.7), compared to an unconditional mean promotion rate of 7.4%. However, after the shock, the coefficient on post is not significant, and therefore contributions to treated firms are not useful for promotion. In contrast, open-source contributions to control firms have a lower effectiveness rate in the full sample, and have a higher rate in the post-period.

Table 9 tests if the effect is stronger for promotion in the same organization where contributions were made (column 2), and if they were useful for promotions in other organizations (column 3). Contributions are more likely to yield a promotion in the same firm (mean of 5.4%) than other firms (1.9%), unconditionally. The variable of interest is $1(Post) \times \log(code_{ext,treat})$; the rate of promotion is lower after the shock for contributions to treated firms. This is consistent with the interpretation that treated firms are financially constrained and therefore are unlikely to hire their contributors. Additionally, in column 3, the rate of promotion to internal developer is similar for both contributions to treated and control firms. I test whether these coefficients are statistically different by first constructing Δ , defined as log contributions to treated minus log contributions to control firms. The null hypothesis is that contributions to either type of firm does not have a differential impact on rate of being hired at a different firm, and therefore the regression coefficient should be zero. Column 3 shows that the coefficient from the *post* interaction term is negative, which suggests that workers are less likely to be hired at a different firm when making contributions to treated firms.²⁴

These results suggest that the marginal volunteer is of worse quality for treated firms

GitHub-LinkedIn profiles because that has a low rate of successful matches and limits the number of observations per firm.
²⁴The difference is negative and statistically significant despite the coefficients in column 3 being similar. This is because the standard errors are larger for the treated group. There are many more control firms, which allows more precision in determining that the control group is statistically different than the treated group. As a robustness check, Appendix Table IA.VIII finds similar results when using shares of a worker's contributions as the explanatory variable, instead of volume of contributions.

after the shock. Even after controlling for observable worker quality and worker fixed effects, contributions to treated firms were less likely to help an external volunteer to be “promoted” to internal status. I interpret this as evidence that, as firms increase their reliance on public projects, public projects include lower quality workers and make it more difficult for high-types to use these projects as a signaling device.

The framework also gives predictions on the types of projects that will be released as open-source projects. These should be projects that would have been closed-source, but now firms release as open-source due to financial constraints. If this is true, then I hypothesize that these projects should be relatively more sophisticated than the projects in the pre-period. To test this, I use a large language model to rate code sophistication.²⁵ The null hypothesis is that there would be no difference in sophistication. In Table [IA.VII](#), I present results and find that sophistication increased by 4.8, which is large compared to a mean of 41 and standard deviation of 21. I interpret this result as suggesting that the new projects were more meaningful to the firm, for example, if they contained more sophisticated intellectual property.

With the sophistication score in hand, I also return to the tests on worker signaling to evaluate whether hiring differences can be accounted for based on code sophistication. In Table [IA.IX](#), I find that the probability of hiring does increase with code sophistication, but the spillover effect still persists even when controlling for this proxy of observable quality.

Additional Evidence for the Signaling Mechanism

The hypotheses embed additional implications for firms and workers that can be tested in the data. I test implications in four domains: firm outcomes, project characteristics, and existing and new employees of shocked firms. First, for firms, a first-order implication is that firms must face some cost associated with open-source development. Otherwise, if open-source projects were always beneficial, then firms should always pursue open-source

²⁵I use OpenAI Chat Completions and ask it to grade code sophistication between 0-100 based on a rubric. I randomly sample 15% of code contributions to be scored.

development even absent the shock.

I assemble a monthly panel of firm deals with dummy variables for if a firm undergoes bankruptcy, IPO, M&A, secondary transactions, or VC funding. The panel is truncated if bankruptcy, IPO, or M&A occur because there would be no further observations afterwards. For the same reason, this table does not use firm fixed effects. The literature typically views bankruptcy as a negative outcome, IPO and M&A as positive outcomes, and Secondary Transactions or VC funding are less unilaterally positive or negative. Table 10, subpanel A presents results from a difference-in-differences specification with the dummy variable multiplied by 100 for scaling. As already discussed, the prevalence of VC funding declines. The rate of IPO also declines: previously, the sample had an unconditional mean of 0.02% of months with an IPO. When a firm is financially constrained, the rate declines by 0.06 percentage points, after accounting for date fixed effects. Subpanel B focuses on observations when a firm reports whether the next VC round was up, flat, or down, based on the change in the valuation. The likelihood of an up round for the constrained firms is much lower. Whereas a firm reports an up round in 1.6% of months unconditionally, the likelihood declines by -2.4%, a 1.5x change, for treated firms after the shock.

The prior tables present evidence from financial constraints but do not speak to the effect of open-source. As a first approximation, I use the presence of an active GitHub account as a proxy for the ability to use open-source. Table IA.XV shows that the channel is stronger for firms with GitHub accounts. Table IA.XVI presents a triple difference-in-difference, and finds that the rate of IPO was lower for those with GitHub accounts. Therefore, while treated firms with GitHub accounts were able to have more headcount growth, they ultimately were less likely to IPO.

Second, I examine whether project characteristics could be consistent with worker signaling incentives, such as if some projects are more relevant as a signaling device than others. I use the primary programming language of a project to determine whether variation in

project characteristics drives more volunteers. Most repositories have one language that is the dominant language, as seen in the high share of the largest language in the histogram in Figure IA.3. Therefore, I test whether the repositories of treated firms have a greater share of volunteer contributions based on language.²⁶ The interpretation is that volunteers would make contributions in more high-value languages, either because it is a rare skill or the skill has more potential employers. I run a triple-difference-in-differences with an additional indicator based on whether the language is in the top 30 languages, with Python as the base indicator variable and “Other” for all other primary languages. I plot all the interaction terms with language dummies in Figure IA.4. I find that these dummies jointly have an F-stat of 1.57, which is not statistically significant at the 1% level. From comparing coefficients, it appears that “Other,” or not being in a top language, attracts slightly more volunteers. Python, which has lower barriers to entry, does not disproportionately attract more volunteers. Rust, Kotlin, HCL attract the most volunteers, as evidenced by the large *t*-statistics. What makes these languages similar is that these are newer, general-purpose languages. I compare the *t*-statistic of each language to the date the language was released and present scatter plot in Figure IA.5, but the statistical relationship is not strong. Overall, this is suggestive evidence that new yet broadly applicable programming languages more easily attract volunteers, possibly because there is a larger future payoff.

Third, I use a panel of job histories of developers and examine the developers that start after the FTX collapse. Table IA.X shows that new hires after the collapse earn -2.3% lower wages, compared to other developers in the same firm before the collapse, and absorb any developer-specific wages. However, there is no change in seniority bin or position number, suggesting that these position characteristics have not changed despite lower corresponding wages.²⁷ This reinforces the idea that firms may have been financially constrained after the shock. However, it may also reflect a trade-off where firms can compensate workers with

²⁶The regression is weighted by volume of contributions, as explained in Table 4, and therefore it would recover the same effect regardless of whether the unit of observation is at the organization-quarter level or the organization-month level.

²⁷Seniority is an integer from 1 to 7 and is a variable created by Revelio Labs based on job title and other position observable characteristics. Position number denotes the job’s index in a worker’s employment history.

lower wages when more projects are public because workers value the public projects. To shed additional light, the data on job histories provides data on worker characteristics. These workers appear 3.2% less likely to be in the U.S., and 9.2% less likely to be in the same state as the modal employee before the shock. This suggests that firms were able to benefit from improving access, which is similar to the spirit of open-source development.

Lastly, I consider whether firms are harmed from open-source development due to an increase in entrepreneurship. Following a theme from [Babina \(2019\)](#) that employees of distressed firms are more likely to enter entrepreneurship, I track the career progression of workers who were employed at the treated firms at the time of the shock and present results in Table [IA.XI](#). After the shock, employees at the treated firms earned -2.9% lower wages in future positions. I find that existing employees are 0.086 percentage points more likely to be founders in the future. This effect is 64% of the unconditional mean of 0.1% of all positions mentioning “founder” in the title, a relatively large increase. Interestingly, these employees are also less likely to use open-source in the future. One potential interpretation is that entering entrepreneurship increases competition for the treated firms.

This set of results brings multiple pieces of evidence that suggests open-source development functions as a signaling mechanism. Firms benefit by accessing cheaper labor, but face potential costs such as possibly facing new entrants. More interestingly, workers also face potential costs in that greater interest in open-source makes it more difficult for high-type workers to signal their ability and separate from other workers.

5 Robustness Tests

This section presents additional analysis and first explores heterogeneity by firm age and size to understand the sources of variation in the data. I then discuss alternative treatment specifications and alternative channels.

5.1 Heterogeneity and Firm Dynamics

I briefly explore whether small firms or young firms are more likely to host more public projects during the 2022 shock. I first provide basic summary statistics on what types of firms are unconditionally more likely to have public projects. I collect data on all VC investments in PitchBook and label whether they have a successful match to a GitHub account. I use the lack of a match as a proxy to suggest that these firms do not host open-source projects. I run a regression of the form

$$\begin{aligned} 1(\text{HasGitHub})_i = & \beta_0 + \sum_{a=1}^A \beta_a \mathbf{1}\{\text{AgeDecile}_i = a\} + \sum_{s=1}^S \beta_s \mathbf{1}\{\text{SizeDecile}_i = s\} \\ & + \sum_{j=1}^J \gamma_j \mathbf{1}\{\text{Industry}_i = j\} + \sum_{k=1}^K \delta_k \mathbf{1}\{\text{Location}_i = k\} + \varepsilon_i. \end{aligned}$$

Formally, with respect to each set of fixed effects, the F-statistics for industry is 248.37, age is 188.44, location is 104.66, and size is 40.72. This suggests that industry is the largest determinant of prevalence of public projects, followed by age. Figure 8 presents firm dynamics. Larger, older firms are more likely to have open-source projects. Further, there is more variation along the age deciles, because age has more explanatory power than size in driving the appearance on GitHub. The scale is monotonically increasing, except for the newest firms as of 2025, which are slightly more likely to be matched on GitHub.

With these findings in hand, I turn to evaluate whether the main results are driven by a lack of matches, especially in small or young firms. I estimate the treatment effect with an additional interaction term to indicate if firms are small or young firms, compared to the median. Table IA.XIV, column 1, shows that the decline in workers was stronger for small firms and young firms, but the treated firms are still robust controlling for these effects. Column 2 similarly shows that small firms exhibit the largest shifts to open-source software after the shock. Young firms, which are more likely to have a GitHub, do not seem disproportionately more likely to attract volunteers. Importantly, firm age drives the

largest variation in matching rates, but does not appear to be associated with differential results, which suggests that the results are less likely to be driven by poor GitHub-PitchBook matching.

5.2 Alternative Treatment Specifications

In this setting, the most prominent endogeneity concern is that the treatment group is unobservably different than the control group due to VC selection effects. It is possible that VCs choose firms that are more inclined to rely on the open-source community. I explore this in three ways: first, I evaluate whether some VCs have some “treatment” effect in encouraging firms to pursue open-source development. Second, a research design to match similar firms into cohorts mitigates selection issues from VCs alone. Lastly, I present four alternative assignments of treatment groups to evaluate the robustness of the main test.

I estimate a panel regression of external share on indicator variables of the top 50 largest VCs. The F -statistic across all indicator variables is 388, which suggests that there is some selection effect between VCs and startups with respect to external share. Figure [IA.2](#) presents the VC t -statistics and highlights the treated VCs in orange. The treatment group is dispersed around zero and none are statistically significant.²⁸ I also estimate a difference-in-differences specification including indicator variables for VCs. Results in Table [IA.XII](#) show that the substitution effect is still strong after including these controls. Altogether, lack of effects partially mitigates the concern that the treatment group is selected based on VCs or firms that are more inclined to use open-source development.

I implement a matching design by forming cohorts for each treated firm. Control firms are selected with replacement based on industry, location, age, and number of GitHub contributors as of 2021. Industry is determined by PitchBook-designated primary industry and location requires firms in the U.S. to be in the same state and firms outside of the U.S. must be in the same country. I require control firms to be within 5 years of the same age, and have

²⁸Note that there is no “base” VC as the indicator variable because multiple VCs may be active in each firm and therefore there is no perfect colinearity concern.

the number of contributors within a 25% bandwidth. This last criterion also requires treated and control firms to have a GitHub presence prior as of 2021. The results are presented in Appendix Table [IA.XIII](#). The coefficient for headcount growth is -4%, which is larger in magnitude than the main result with controls, while the coefficient on external share is smaller at 8.5%. Matched sample results are qualitatively similar to the main specification.

It may be possible that this is driven only by a subset of VC firms. The following three tables present four robustness tests from exploring the heterogeneity within the treatment group and a placebo group. First, Table [IA.XVIII](#) splits the treatment group into two sets: treated firms that are associated through large VC firms, or treated firms that are associated with all other VC firms. The first set is tested in column 1 and 3, while the second is in column 2 and 4. The names of the selected VCs are in the table caption. The results appear to have economically similar coefficients for both sets, but results are more precise in the small VCs. I similarly test whether the results are strong when isolating the large crypto-focused VC firms. Similar to the prior table, Table [IA.XIX](#) displays results in these narrower sets of treated groups. The substitution effect is similar to the main specification, where headcount growth is negative and external work is positive, but crypto-focused VCs saw more dramatic declines in headcount growth and a stronger increase in external share of work.

As a placebo test, I consider whether other large VCs saw a similar result. Table [IA.XX](#) displays results for two placebo tests. I first pick five well-known VC firms as a treated group. The results suggest that while these companies also appear to grow slower, they did not see an increase in open-source development. I then suppose high-volume VCs were the treated group. Here, the data also shows a decline in headcount growth, but not a shift into open-source development. However, based on the standard errors, we can see that the decline in headcount growth is smaller than those that are related to FTX. Thus, the shock to VCs seemed to most strongly impact those that invested in FTX, and particularly the

smaller VCs, regardless of crypto-focus.

5.3 Alternative Channels

There are other competing channels that may rationalize my results. One alternative narrative is that firms may have both public and private projects, and that constrained firms may focus on more critical projects privately and allow non-critical public projects to languish. I cannot observe the total share within a firm of projects that are public or private; however, these results from Table 5 suggest that firms are actively growing the public portion. The private portion is likely not growing as fast because the prior results show that the number of workers is not increasing as quickly, and by a resource constraint argument, it is likely that the private projects are not growing as quickly as the public projects. The opposite directions of these effects suggest that the share of projects that are public is likely to increase.

Another alternative channel is that firms prefer open-source because it allows firms to evaluate developers prior to full time employment. The previous results on signal effectiveness already speak to this channel. Table 8 suggests that being “hired” is much more likely at the same company where a worker made contributions. However this effect disappears for treated firms after the shock, and the net result is that treated firms were less likely to hire than control firms, even after controlling for observable worker quality.

Firms may also post more projects publicly in an effort to prepare to raise more funding, as in [Conti et al. \(2025\)](#). This mechanism does not directly conflict with the wage-signal substitution channel. It is indeed possible that firms post more projects online as advertising. However, within the context of the labor substitution channel, this may be a substitute for hiring more sales executives. Posting projects publicly also lowers the barriers for potential customers to try your product.

In sum, this section has evaluated the alternative tests and competing channels to explain results. The empirical results remain consistent with the idea that firms substitute away from

hiring and towards relying on the open-source community. This channel is not isolated in one subset of firms nor VCs. I also test additional implications for open-source, including whether firms shifted project allocation in a way that would be inconsistent with the hypotheses, and if open-source is used as advertising.

6 Equilibrium Model

I develop a model with the goal of formalizing the paper’s hypotheses and rationalizing the results on how average worker quality declines if more firms use public projects. Let there be two competitive labor markets in which firms and heterogeneous workers endogenously select into markets. There is a continuum of both firms and workers that enter this market.

There is a continuum of firms indexed by j with total mass of 1. Firms have production functions that produce cashflow y as

$$y = \alpha_s(\theta l)^\lambda$$

where θ is worker quality. Firms maximize the profit function

$$\max_{s \in \{\text{public}, \text{private}\}, L} \alpha_s E[\theta^\lambda | s] l^\lambda - wl$$

subject to a working capital constraint

$$wl \leq A_j$$

where $E[\theta^\lambda | s]$ will depend on the quality of workers in each market s . I focus on the case with decreasing returns to scale: $\lambda \in (0, 1)$.²⁹

There are two types of firms. A mass z of firms exogenously have a working capital constraint $A_j = \bar{A}$ that requires $wl < \bar{A}$. The mass $1 - z$ of firms have $A_j = A^\infty$ that will

²⁹Without decreasing returns to scale, then generally all labor will be allocated to the more productive technology. Similarly, if public output is more productive than the private output, and given the payoffs to workers for public output, then all firms will pursue public output.

never bind.

Firms choose whether the mode of work for a given project is private or public. I normalize $\alpha_{private} = 1$ for all private projects. If a firm chooses a public project, profits will be reduced to $\alpha_{public} = \delta < 1$. This discount captures that profits are reduced because other competitors can copy this project by introducing a substitute.³⁰

There are two competitive labor markets indexed by s : public and private. Define expected labor augmentation as

$$g_s = E[\theta^\lambda | s]$$

From the firm FOC, labor demand for public work as a function of wages is:

$$l_j = \left(\frac{w_{public}}{\delta \lambda g_{public}} \right)^{\frac{1}{\lambda-1}} \quad (2)$$

Define aggregate labor demand for public work as:

$$L_{d,public}(w_{public}) = \int_{j \in J} l_j(w) dj \quad (3)$$

$$= \left(\frac{w_{public}}{\delta \lambda g_{public}} \right)^{\frac{1}{\lambda-1}} \quad (4)$$

where the domain of the integral assumes that all firms have demand for public work. Similarly, private work demand is

$$L_{d,private}(w_{private}) = \left(\frac{w_{private}}{\lambda g_{private}} \right)^{\frac{1}{\lambda-1}} \quad (5)$$

There is a unit mass of workers. They maximize the payoff function

$$\max_{s \in \{public, private\}} w_s + \mathbf{1}_{s=public} R(\theta_i, \theta_{-i})$$

³⁰If some firms had $\delta > 1$, they earn a premium when pursuing a public project, which captures the idea that firms pursue public, open-source projects because it is a complement to their other lines of business. For example, Google releases browser software to support their search business. Firms with $\delta > 1$ would be more likely to choose a public project.

subject to an outside option that requires $u \geq 0$, and limited liability such that $w \geq 0$. To facilitate these markets, I assume that workers only supply indivisible labor to one market.

Workers that participate in the market for public work will receive a benefit $R(\theta_i, \theta_{-i})$ for the worker θ_i . This captures the idea that public work is more visible and increases the expected present value of future wages. $R(\theta_i, \theta_{-i})$ is increasing in θ_i , because higher types that participate in public work will increase the expected type of worker that participates in public modes of work. I assume that workers in the private market have R benefit normalized to zero.

An equilibrium will be wages $w_{private}, w_{public}$, choices s_i for all workers, and choices s_j for firms of each type, such that

1. Firms of types $\bar{A}$ and A^∞ each choose s_j to maximize profits, hires labor L
2. Workers of types θ_i each choose s_i maximize payoffs
3. Labor markets clear for both private and public work

If a cutoff equilibrium for workers exists, there will be θ^* that is indifferent between public and private work. Further, any $\theta_i > \theta^*$ prefers public wages and R while $\theta_i < \theta^*$ prefers private work. It must be that $\theta_i > \theta^*$ prefers public (as opposed to private) because $R(\theta_i, \theta_{-i})$ is increasing in θ_i . One natural choice of $R(\theta^*, \theta_i)$ would be if workers were paid as their expected type:

$$R(\theta_i, \theta_{-i}) = E[\theta_i | \theta_i > \theta^*]$$

The cutoff equilibrium will be θ^* such that

$$w_{public} + R(\theta^*, \theta_{-i}) = w_{private}$$

$$R(\theta^*, \theta_{-i}) = w_{private} - w_{public}$$

There will be a unique cutoff if the left-hand side and right-hand side have derivatives with

respect to θ^* of opposite signs. If more workers want to do private work, then θ^* increases, w_{public} increases and $w_{private}$ decreases, so $w_{private}(\theta^*) - w_{public}(\theta^*)$ decreases. $R(\theta^*, \theta_{-i})$ is increasing with θ^* . Therefore, there is a unique cut-off that pins down θ^* based on the parameterization of $R(\theta^*, \theta_i)$. In equilibrium, we also see $w_{private} > w_{public}$ as a strict inequality.

Because workers supply indivisible labor, then labor supply for each type of work will be a mass of workers such that

$$L_{s,private} = F(\theta^*) \quad (6)$$

$$L_{s,public} = 1 - F(\theta^*) \quad (7)$$

where $F(\cdot)$ is the CDF of type θ .

With respect to firms, I check the following cases of equilibrium:

1. $\bar{A}$ is such that $w_{public}l_{public} < \bar{A} < w_{private}l_{private}$, and therefore z mass of firms only participate in public markets
2. $\bar{A}$ is sufficiently large such that $wl < \bar{A}$ never binds, and therefore both types of firms participate in both markets

This section focuses on characterizing the separating equilibrium, and other cases are relegated to the appendix. A separating equilibrium among firms would require that the mass z of constrained firms chooses the option with lower wages, the mass $1 - z$ of unconstrained firms choose the option with higher wages, and neither type of firm has a profitable deviation.

In that setting, then labor demand would be segmented:

$$L_{d,public}^z(w_{public}^z) = \int_z l_j(w) dj = z \left(\frac{w_{public}}{\delta \lambda g_{public}} \right)^{\frac{1}{\lambda-1}} \quad (8)$$

$$L_{d,private}^z(w_{private}^z) = (1 - z) \left(\frac{w_{private}}{\lambda g_{private}} \right)^{\frac{1}{\lambda-1}} \quad (9)$$

Market clearing requires:

$$z \left(\frac{w_{public}^z}{\delta \lambda g_{public}} \right)^{\frac{1}{\lambda-1}} = 1 - F(\theta^{*z}) \quad (10)$$

$$\Rightarrow w_{public}^z = \delta \lambda g_{public} \left(\frac{1 - F(\theta^{*z})}{z} \right)^{\lambda-1} \quad (11)$$

$$\Rightarrow w_{private}^z = \lambda g_{private} \left(\frac{F(\theta^{*z})}{1 - z} \right)^{\lambda-1} \quad (12)$$

The new indifferent worker must satisfy

$$R(\theta^{*z}) = w_{private}^z(\theta^{*z}) - w_{public}^z(\theta^{*z})$$

The new cutoff θ^{*z} may be smaller or larger than θ^* depending on z . Focusing just on θ^{*z} , the derivative of the right hand side with respect to z is negative, and thus, a larger z yields a smaller $R(\theta)$, which corresponds to a smaller θ^* .

Lastly, the unconstrained firms with mass $1 - z$ must not have a profitable deviation to enter the market for public work. Again using optimal profits as $\pi_j^* = \frac{1-\lambda}{\lambda} w_j^* L_j^*$, and comparing if $\pi_{private}^z > \pi_{public}^z$, equilibrium requires:

$$\begin{aligned} w_{private}^{*z} L_{private}^{*z} &> w_{public}^{*z} L_{public}^{*z} \\ \lambda g_{private} \left(\frac{F(\theta^{*z})}{1 - z} \right)^{\lambda-1} F(\theta^{*z}) &> \delta \lambda g_{public} \left(\frac{1 - F(\theta^{*z})}{z} \right)^{\lambda-1} (1 - F(\theta^{*z})) \\ \frac{1}{\delta} &> \left(\frac{g_{public}}{g_{private}} \right) \left(\frac{1 - z}{z} \right)^{\lambda-1} \left(\frac{1 - F(\theta^{*z})}{F(\theta^{*z})} \right)^{\lambda} \end{aligned}$$

I summarize the separating equilibrium in Figure 9. Panel A summarizes the parameter space where separating equilibrium will exist. The area in green is where it is incentive compatible for unconstrained firms to limit themselves to private markets, and thus will not deviate to public markets because it is not profitable. The separating equilibrium exists

when δ is small, which is the case when the discount from public projects is large.

Panel B shows how the cutoff type θ^* varies in the parameter space δ, z . Along the horizontal axis, if the profitability discount from public projects is large ($\delta \ll 1$) then firms will pursue private projects, demand will be low for public projects, and thus wages will be low, yielding only the high-type workers will participate in open-source. On the vertical axis, if more firms are financially constrained, then the cutoff declines, which suggests that the marginal worker is of lower quality. This rationalizes the main empirical findings.

7 Conclusion

This paper investigates how financially constrained firms can substitute between hiring more labor and relying on volunteer developers, and whether this has externalities for resolving asymmetric information. Treated firms increase their reliance on external developers and post more projects on GitHub. However, the marginal volunteer appears to be of lower quality, and contributions to treated firms have a weaker relationship with future hiring after the shock. Altogether, these results demonstrate strong linkages between incentives of firms and the open-source community.

The mechanism in this paper offers an explanation for why open-source development is so popular. Workers prefer open-source development because it is a device for Spence signaling. High-quality workers can exert costly effort to make a volunteer contribution in return for higher future wages. In addition to the obvious benefit that unpaid labor is cheaper than paid labor, firms also can use open-source to induce separation, which decreases future search costs in finding and hiring high-quality developers. Moreover, the cost of mitigating asymmetric information is ultimately borne by the developers who forgo wages. Firms do not internalize the externalities from a potential over-reliance on open-source developers, such as if it reduces worker signaling incentives.

Given the large and growing role of open-source software, future researchers should continue to explore the complex interaction between firm and worker incentives. One particu-

larly important topic is how sharing intellectual property impacts industry dynamics. For example, workers may benefit from when there is more coordinated open-source software because it makes their work more visible and their human capital more portable. Despite the clear preferences of workers, there is wide variation in how quickly industries have adopted open-source. Prominent early examples include Netscape, a publicly traded company at the time, which released its source code at the height of the browser wars in 1998. In contrast, NVIDIA, which is the world's most valuable company at the time of this writing, has only released its GPU kernel as of 2022, almost 24 years later. Within industry dynamics, it is an important empirical question whether open-source undercuts their smaller competitors, or if it spurs more growth by reducing barriers to entry and mitigating information asymmetry among workers.

References

- About El-Komboz, Lena, and Moritz Goldbeck, 2024, Career concerns as public good-the role of signaling for open source software development .
- Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt, 2005, Competition and innovation: an inverted-u relationship, *The Quarterly Journal of Economics* 120, 701–728.
- Babina, Tania, 2019, Destructive creation at work: How financial distress spurs entrepreneurship, *The Review of Financial Studies* 33, 4061–4101.
- Bernstein, Shai, Kunal Mehta, Richard Townsend, and Ting Xu, 2022, Do startups benefit from their investors’ reputation? evidence from a randomized field experiment, *SSRN Electronic Journal* .
- Boysel, Sam, Manuel Hoffmann, and Frank Nagle, 2024, Labor competition and open innovation, *Working Paper* .
- Campello, Murillo, John R. Graham, and Campbell R. Harvey, 2010, The real effects of financial constraints: Evidence from a financial crisis, *Journal of Financial Economics* 97, 470–487, The 2007-8 financial crisis: Lessons from corporate finance.
- Chen, Jun, and Michael Ewens, 2025, Venture capital and startup agglomeration, *The Journal of Finance* 80, 2153–2198.
- Chodorow-Reich, Gabriel, 2013, The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis, *The Quarterly Journal of Economics* 129, 1–59.
- Coleman, Braiden, Karson Fronk, and Kristen Valentine, 2025, Corporate adoption of an open innovation strategy: Evidence from github .
- Conti, Annamaria, Christian Peukert, and Maria Roche, 2025, Beefing it up for your investor? engagement with open source communities, innovation, and startup funding: Evidence from github, *Organization Science* 36, 1551–1573.
- Eisfeldt, Andrea L., and Adriano A. Rampini, 2007, New or used? investment with credit constraints, *Journal of Monetary Economics* 54, 2656–2681.

- Eisfeldt, Andrea L., and Adriano A. Rampini, 2008, Leasing, ability to repossess, and debt capacity, *Review of Financial Studies* 22, 1621–1657.
- Eldar, Ofer, and Jillian Grennan, 2023, Common venture capital investors and startup growth, *The Review of Financial Studies* 37, 549–590.
- Emery, Logan, Chan Lim, and Shiwei Ye, 2025, The private value of open-source innovation, *SSRN* .
- Ersahin, Nuri, Ruidi Huang, and Naveen Khanna, 2021, The achilles heel of reputable vcs, *SSRN Electronic Journal* .
- Garmaise, Mark J., 2007, Production in entrepreneurial firms: The effects of financial constraints on labor and capital, *Review of Financial Studies* 21, 543–577.
- Gortmaker, Jeff, 2025, Open source software policy in industry equilibrium .
- Gu, Lifeng, Ruidi Huang, Yifei Mao, and Xuan Tian, 2017, How does human capital matter? evidence from venture capital, *SSRN Electronic Journal* .
- Hakim Orman, Wafa, 2008, Giving it away for free? the nature of job-market signaling by open-source software developers, *The B.E. Journal of Economic Analysis and Policy* 8.
- Hoffmann, Manuel, Frank Nagle, and Yanuo Zhou, 2024, The value of open source software, *Harvard Business School Working Paper* .
- Holmström, Bengt, 1999, Managerial incentive problems: A dynamic perspective, *Review of Economic Studies* 66, 169–182.
- Howell, Sabrina, Josh Lerner, Ramana Nanda, and Richard Townsend, 2020, *How Resilient is Venture-Backed Innovation? Evidence from Four Decades of U.S. Patenting*.
- Ivashina, Victoria, and David Scharfstein, 2010, Bank lending during the financial crisis of 2008, *Journal of Financial Economics* 97, 319–338.
- Ke, T. Tony, Christopher Li, and Mikhail Safronov, 2023, Learning by choosing: Career concerns with observable actions, *American Economic Journal: Microeconomics* 15, 536–567.
- Kumar, Vineet, Brett R. Gordon, and Kannan Srinivasan, 2011, Competitive strategy for open source software, *Marketing Science* 30, 1066–1078.

- Lakhani, Karim R, and Eric von Hippel, 2003, How open source software works: “free” user-to-user assistance, *Research Policy* 32, 923–943.
- Leppämäki, Mikko, and Mikko Mustonen, 2009, Skill signalling with product market externality, *The Economic Journal* 119, 1130–1142.
- Lerner, Josh, and Jean Tirole, 2002, Some simple economics of open source, *The Journal of Industrial Economics* 50, 197–234.
- Lerner, Josh, and Jean Tirole, 2005, The economics of technology sharing: Open source and beyond, *Journal of Economic Perspectives* 19, 99–120.
- Llanes, Gastón, and Ramiro de Elejalde, 2013, Industry equilibrium with open-source and proprietary firms, *International Journal of Industrial Organization* 31, 36–49.
- Matsa, David A., 2018, Capital structure and a firm’s workforce, *Annual Review of Financial Economics* 10, 387–412.
- Michelacci, Claudio, and Vincenzo Quadrini, 2009, Financial markets and wages, *Review of Economic Studies* 76, 795–827.
- Nagle, Frank, 2019, Open source software and firm productivity, *Management Science* 65, 1191–1215.
- Nanda, Ramana, and Matthew Rhodes-Kropf, 2013, Investment cycles and startup innovation, *Journal of Financial Economics* 110, 403–418.
- Petralia, Sergio, 2025, Open source software as digital platforms to innovate, *Journal of Economic Behavior and Organization* 237, 107109.
- Roberts, Jeffrey A., Il-Horn Hann, and Sandra A. Slaughter, 2006, Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the apache projects, *Management Science* 52, 984–999.
- Sharma, Vasundhara, Ashish Agarwal, and Anitesh Barua, 2024, Demand-side effects of open innovation: The case of cryptocurrency forking, *Management Science* .
- Spence, Michael, 1973, Job market signaling, *The Quarterly Journal of Economics* 87, 355.

- Sun, Qi, and Mindy Z. Xiaolan, 2019, Financing intangible capital, *Journal of Financial Economics* 133, 564–588.
- Titman, Sheridan, 1984, The effect of capital structure on a firm’s liquidation decision, *Journal of Financial Economics* 13, 137–151.
- Townsend, Richard R., 2015, Propagation of financial shocks: The case of venture capital, *Management Science* 61, 2782–2802.
- Zandberg, Jonathan, Osnat Zohar, Maya Haran Rosen, and Alexander Montag, 2024, Temporal dynamics of venture capital funds: Investment timing and fund performance, *SSRN Electronic Journal* .

8 Figures

Figure 1. Firm share of open-source code by external contributors

This figure plots the share of open-source code written by external contributors on the y-axis, and the size of open-source codebases on the x-axis. This figure only shows organizations with more than 100 code changes, yielding a sample of 613,000 organizations. The figure selectively highlights the code bases of a few very large firms in blue. For the highlighted firms, I also indicate the number of followers accumulated by each firm in the size of the circle.

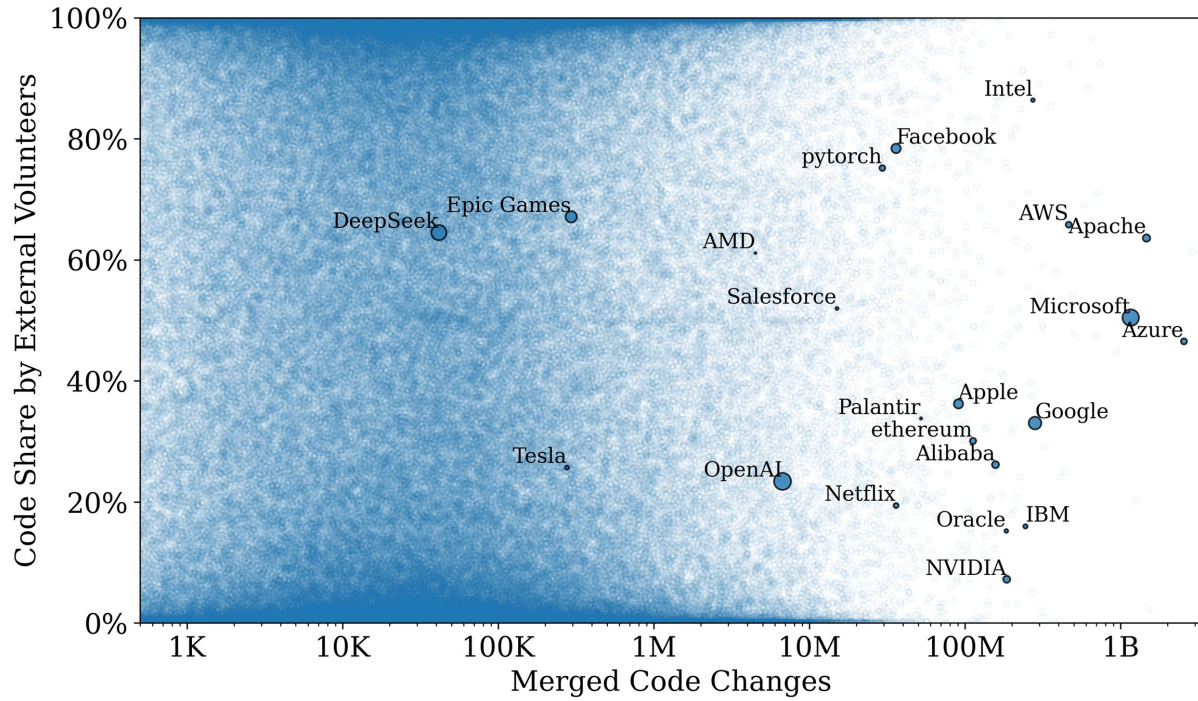


Figure 2. Developer Status: Aggregate Activity and Merge Rates

This figure describes the work contributed by five different types of contributors: Owner, Member, Collaborator, Contributor, and No Status. The left-hand panel includes three bars to describe the share of work by each type. The number of pull requests (PRs) is the share of the count of work submitted by each type. Code changes are the sum of the lines of code additions and code deletions. Number of contributors is the unique number of developers of each type. The right-hand panel is the rate of which these are merged based on the number of pull requests. Orange indicates types I classify as external developers.

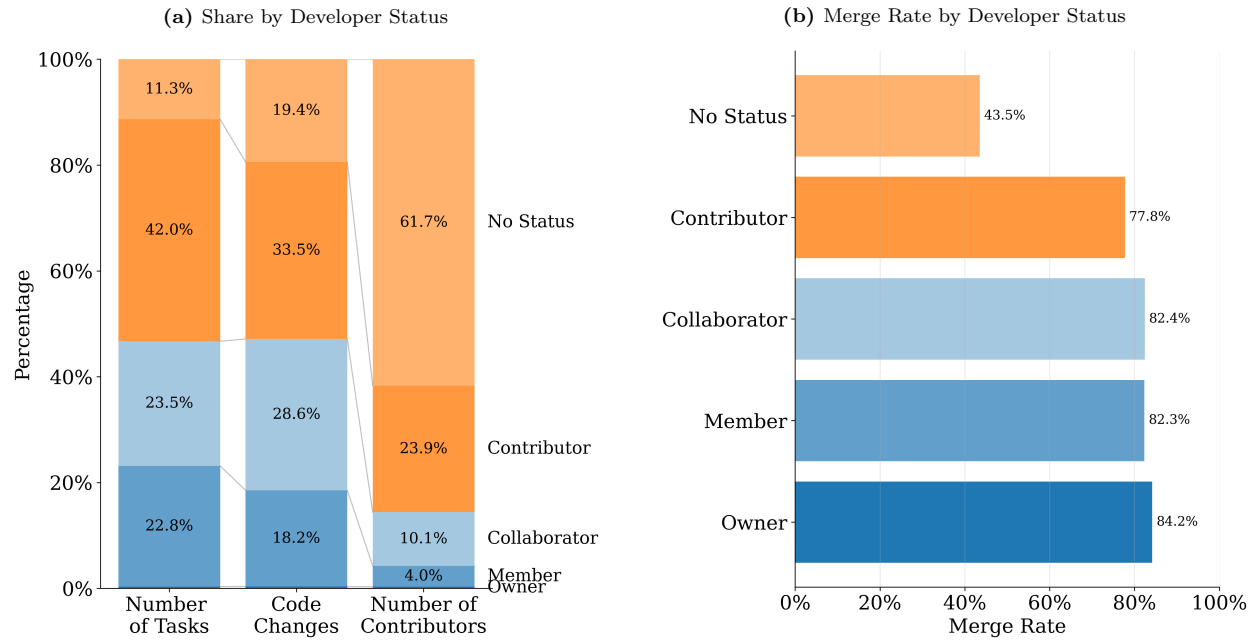
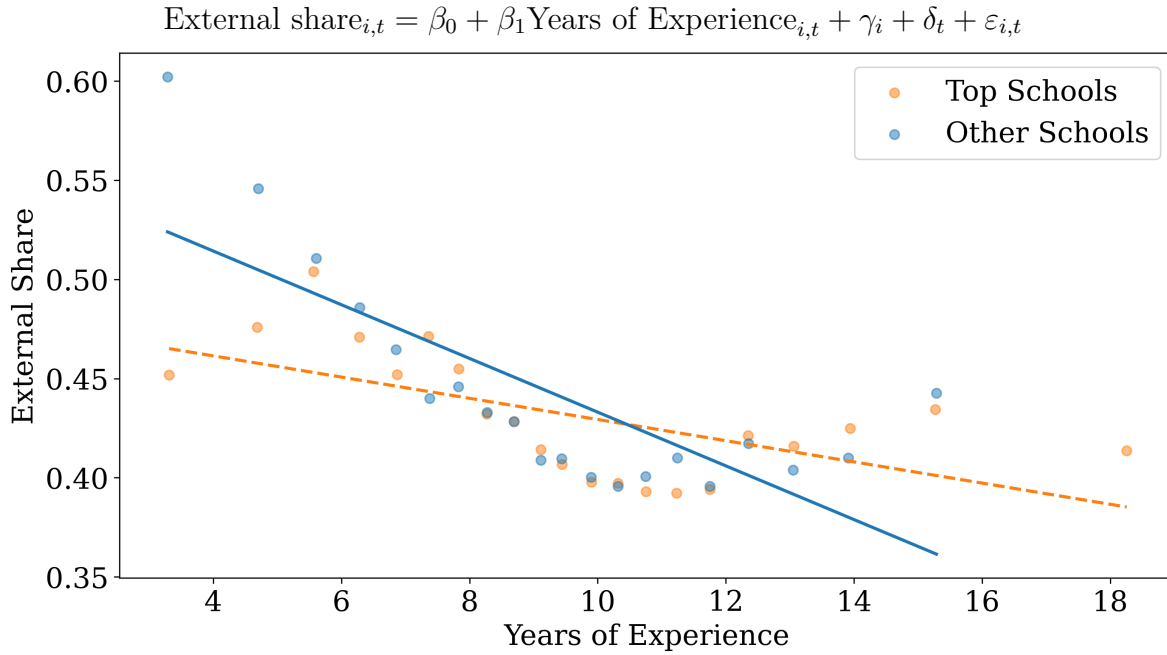


Figure 3. Binscatter: open-source contributions, years of experience, and future wages

This plot shows the relationships share of work that was contributed as an external developer with years of experience (Panel A) and future wages (Panel B). In Panel A, the unit of observation in the panel is developer-year. The x-variable is years of experience based on date of graduation and y-variable is external contributions in that year. The binscatter includes fixed effects by year and developer. In Panel B, the unit of in the panel is developer-position. The binscatter includes fixed effects for current company, departure month, seniority, and location (U.S. states, or country if international). The x-variable is external contributions in that position and the y-variable is the log of total compensation, as predicted by Revelio Labs. Related tables in parentheses.

(a) Open-source contributions and years of experience (Table IA.II)



(b) Future wages and open-source contributions (Table 2)

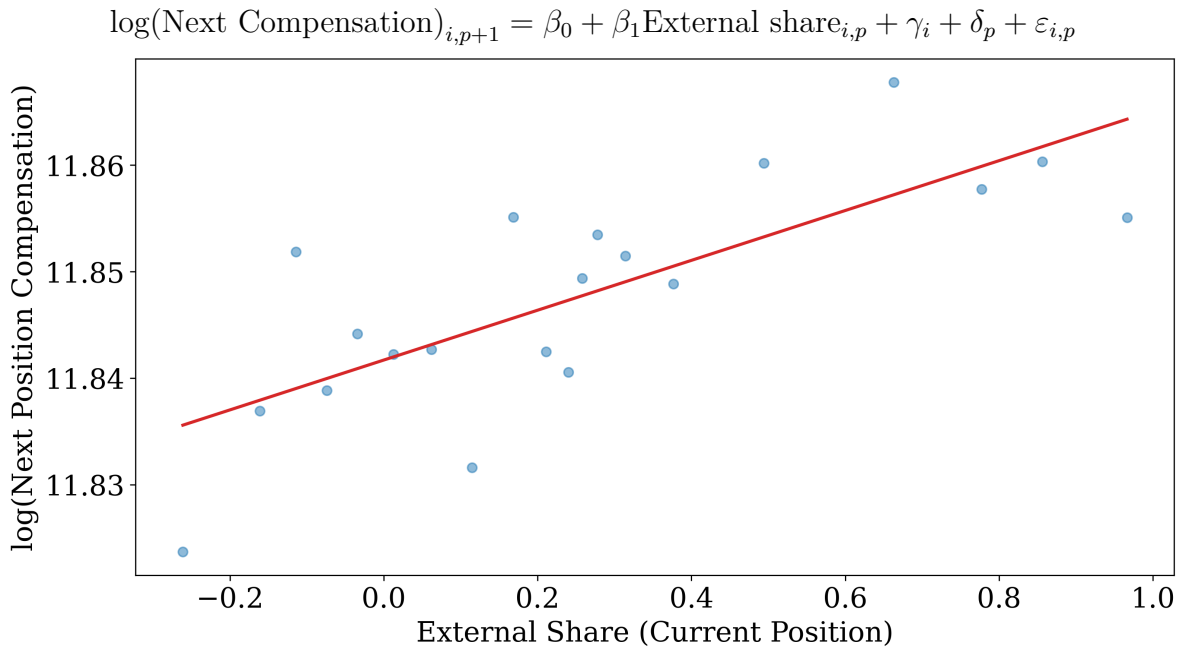


Figure 4. Aggregate volume of accepted code

This figure presents aggregate statistics on the volume of pull requests over time. The top panel plots the monthly volume of bar in light blue bars and the share of work contributed by external developers in lines and dots. The orange line is the treated group and the graph line is the control group. The red vertical lines demarcate the Luna, Celsius, and FTX collapse respectively.

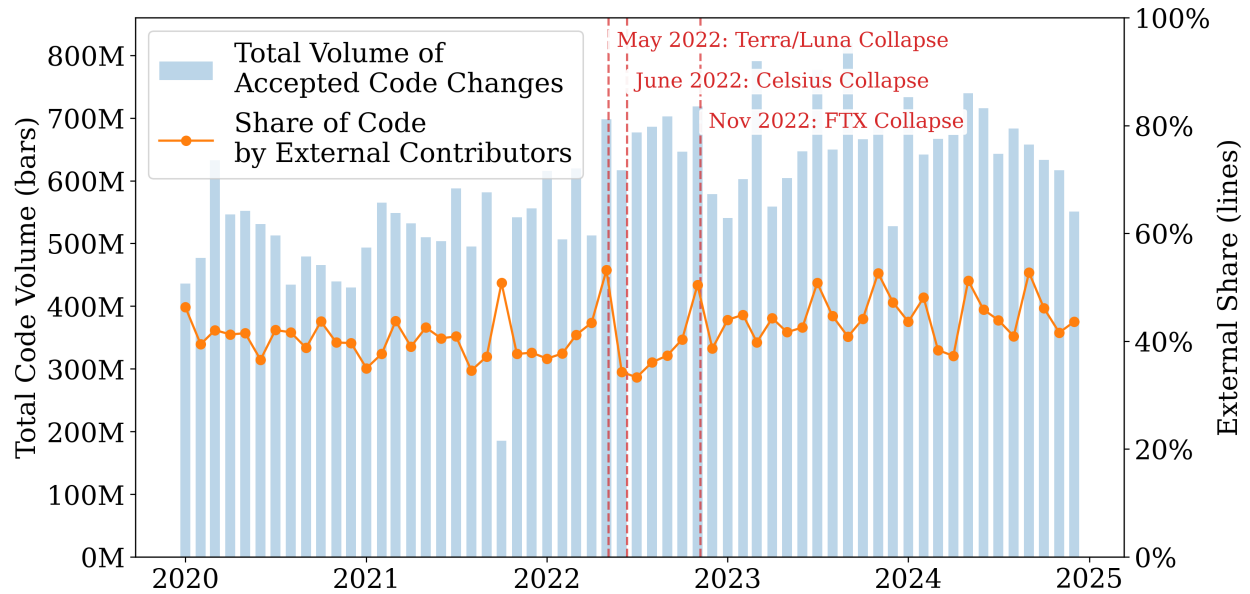
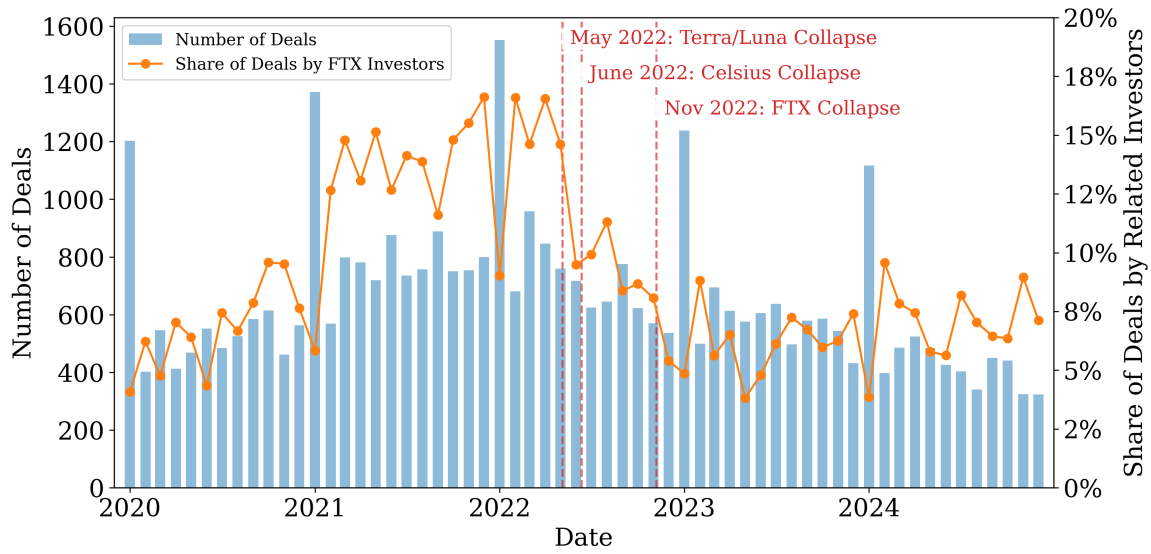


Figure 5. FTX investors and crypto market trends. Panel (a) plots the number of deals completed, split by whether each deal also includes an investor that also invested in FTX. Investors that invested in FTX are labeled “Related” investors while all other investors are “Unrelated.” The left axis plots the bars of number of deals, with orange segments indicating the subset that is by related investors and blue indicating all other deals. Share of related deals, as a fraction of all deals, is plotted on the right hand side with an orange line. The red vertical lines demarcate the Luna, Celsius, and FTX collapse respectively. Panel (b) shows the market capitalization of Bitcoin and the FTX token over time. Bitcoin (BTC) is in blue solid lines, FTX Token (FTT) is in purple dashed lines, and the S&P 500 index is added for comparison. Crypto data is from Coinmarketcap.com, and S&P 500 is from FRED. The three dashed lines show when Terra/Luna, Celsius, and FTX collapsed. The corresponding decline in Bitcoin prices is shown on each day in parentheses.

(a) FTX Investors and Number of Deals



(b)

(c) Bitcoin and FTT Market Cap, and S&P 500 Index

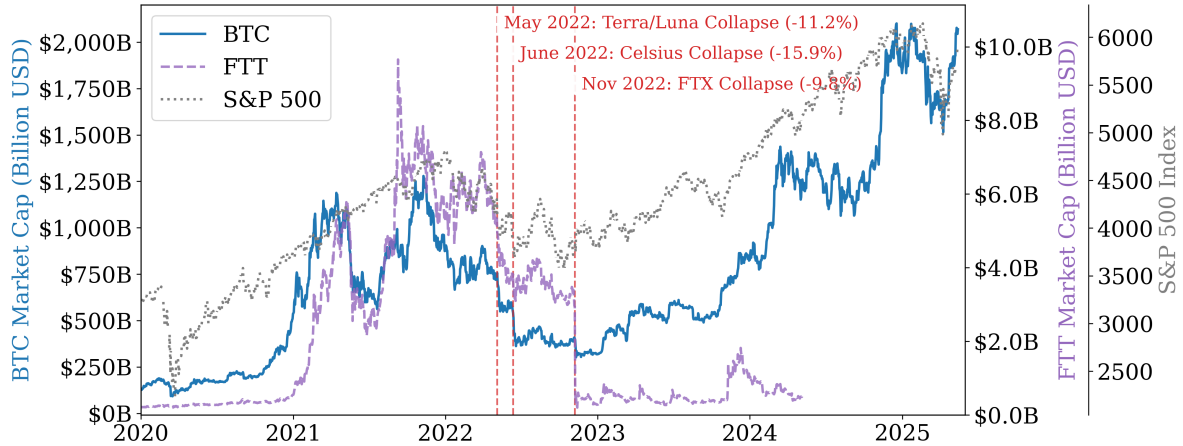
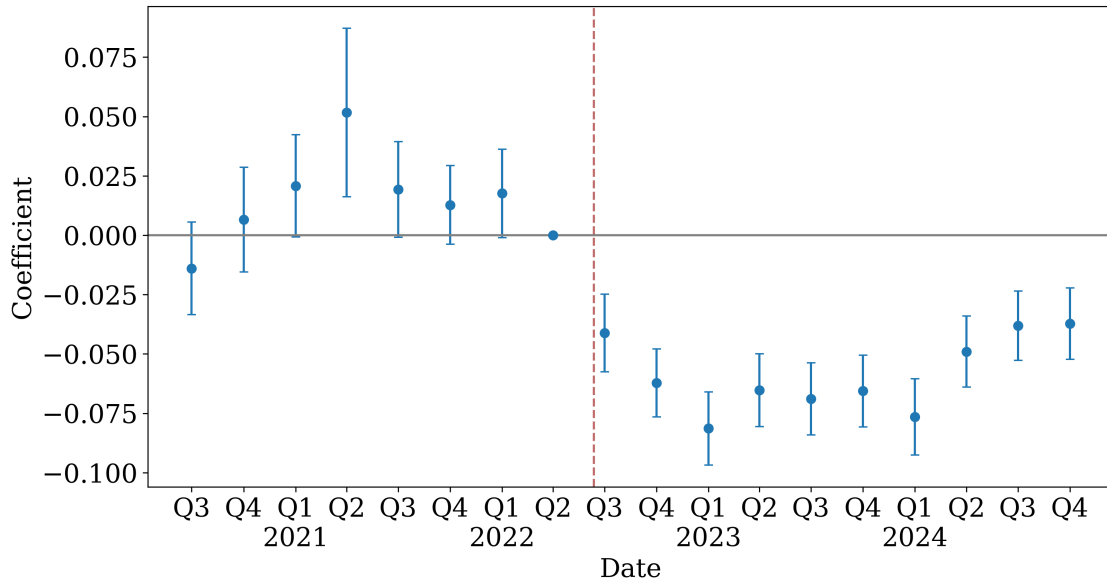


Figure 6. Difference-in-differences: dynamic coefficients

This figure plots the coefficients from a difference-in-differences design with the model below. Panel (a) shows the coefficients for % change in employee headcount and Panel (b) shows share of code by external contributors. Share of code is calculated as the sum of all lines of code additions and deletions by contributors with the status “none” or “contributor” (and thus excludes “collaborators,” “members,” and “owners”) divided by the sum of all lines of code additions and deletions. Error bands reflect 95% confidence intervals. The red dotted line demarcates June 2022, which is when treated VCs began to participate in fewer deals.

$$y_{it} = \beta_0 + \sum_t \beta_t (\text{date}_t \times \text{treat}_i) + \gamma_i + \delta_t + \varepsilon_{i,t}$$

(a) Outcome: employee growth (%)



(b) Outcome: share of code by external contributors (%)

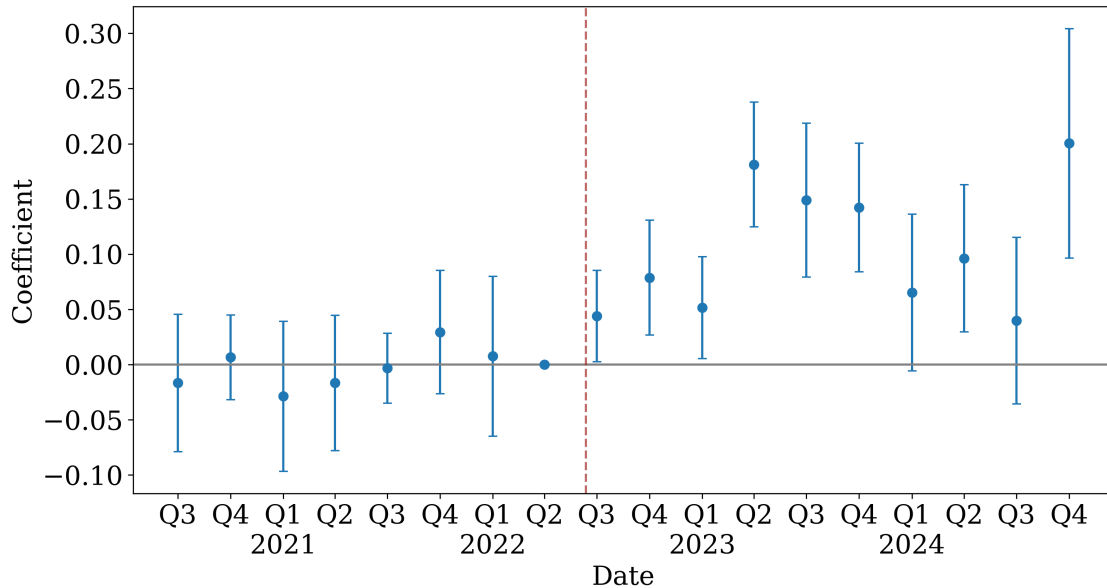


Figure 7. Change in developer merge rates

This figure plots the quality of developers over time. I calculate quality as “rolling changes merge rate” or the share of all code changes (sum of additions and deletions) that have been accepted as a contribution through the lifetime of each developer as of each month. The blue line takes the average of all firms that has an overlapping a financing relationship as FTX and the orange line as all other firms.

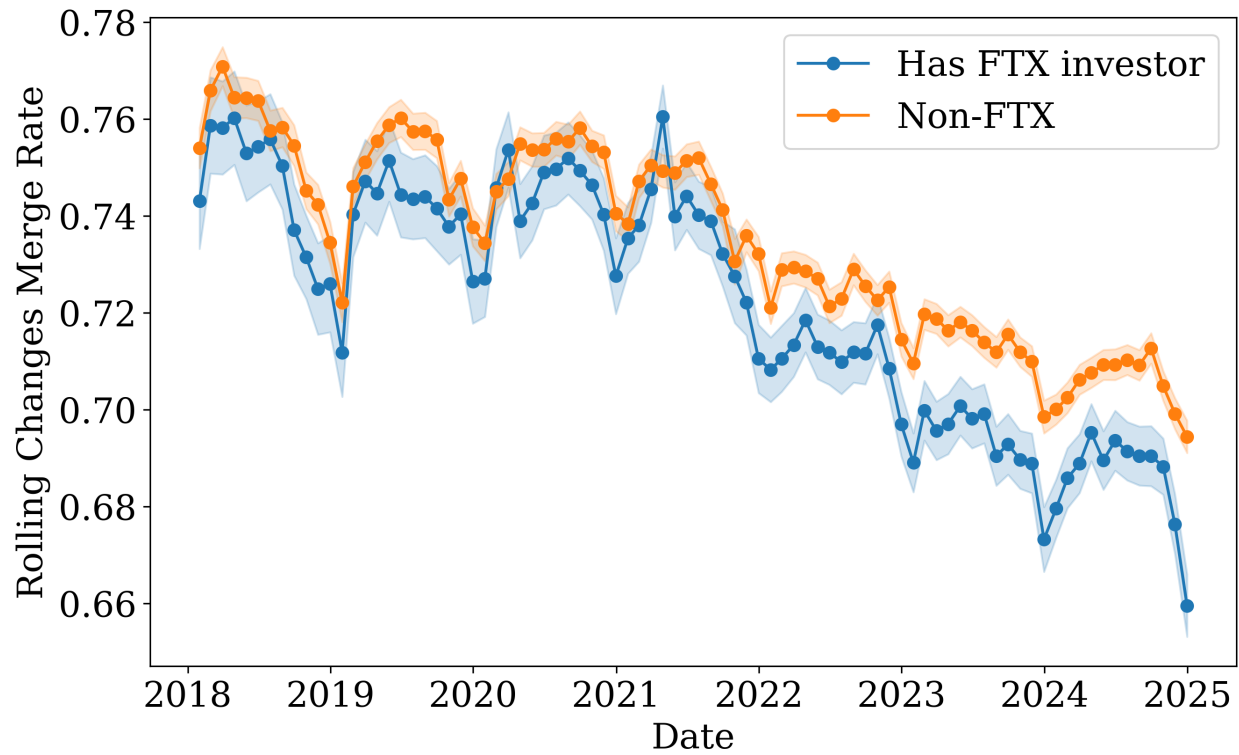


Figure 8. Average rate of matched GitHub by size-age distribution

This graph plots the share of firms that have GitHub accounts, a determined based on whether a match is made between GitHub and PitchBook. Firms are fuzzy-matched based on name, location, and website. Firms are cross-sorted into size deciles by number of employees and age deciles. Decile 10 is the largest. I run the regression below with fixed effects for industry and location to purge the variation driven by these. This figure plots the predicted based on $\beta_0 + \beta_a + \beta_s$, which may yield negative numbers in this specification.

$$\begin{aligned} 1(\text{Has GitHub})_i = & \beta_0 + \sum_{a=1}^A \beta_a \mathbf{1}\{\text{AgeDecile}_i = a\} + \sum_{s=1}^S \beta_s \mathbf{1}\{\text{SizeDecile}_i = s\} \\ & + \sum_{j=1}^J \gamma_j \mathbf{1}\{\text{Industry}_i = j\} + \sum_{k=1}^K \delta_k \mathbf{1}\{\text{Location}_i = k\} + \varepsilon_i \end{aligned}$$

Share of firms with a GitHub presence, by size and age bin

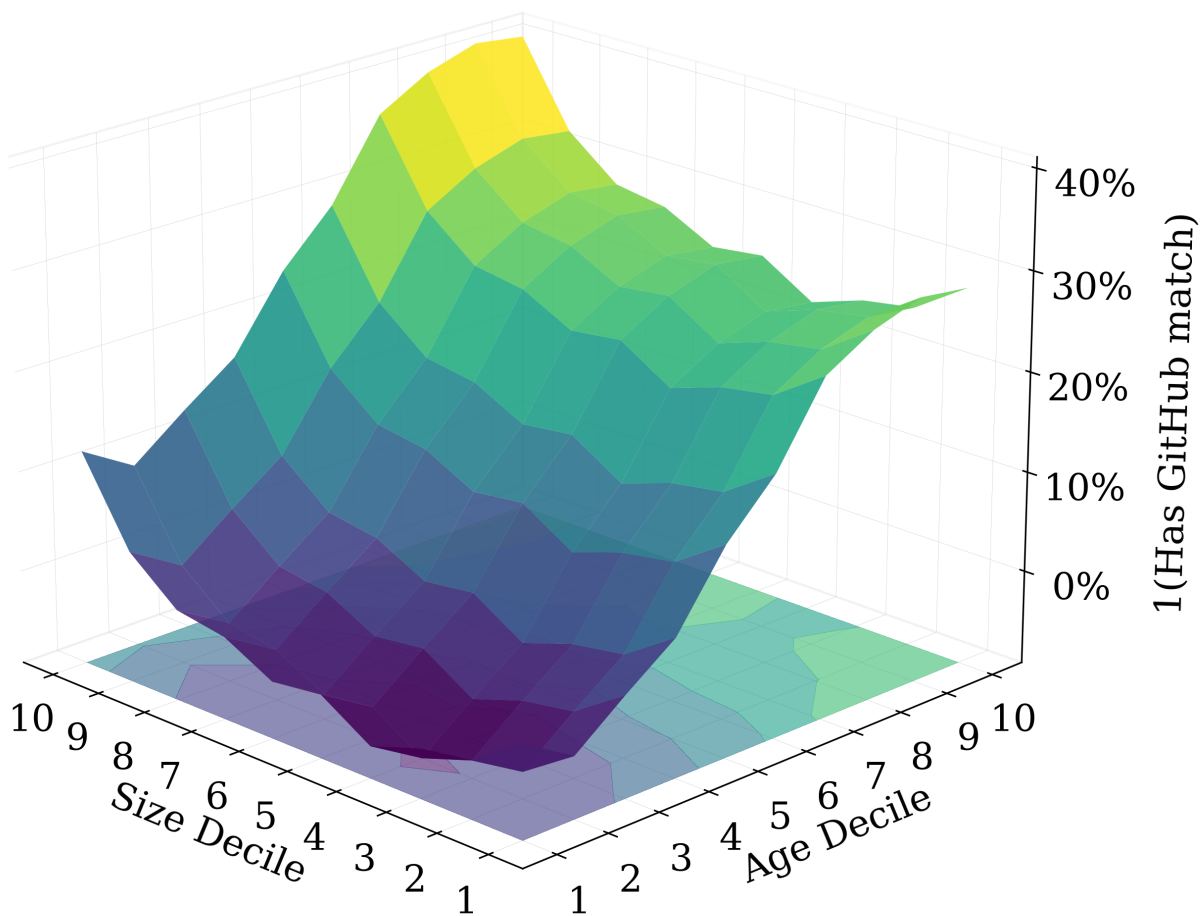
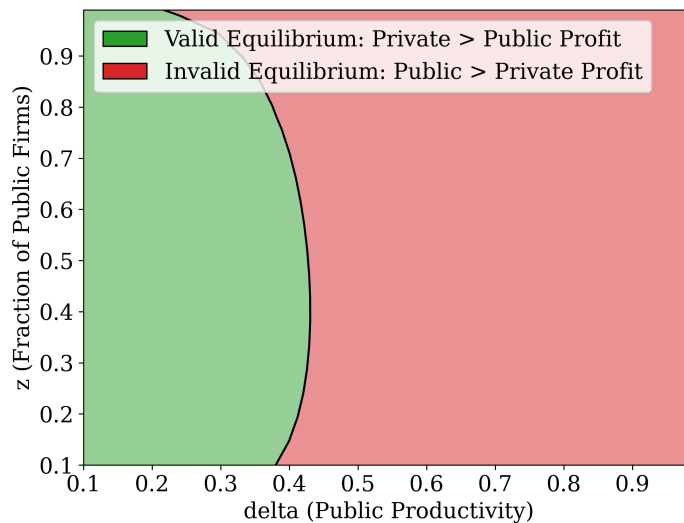
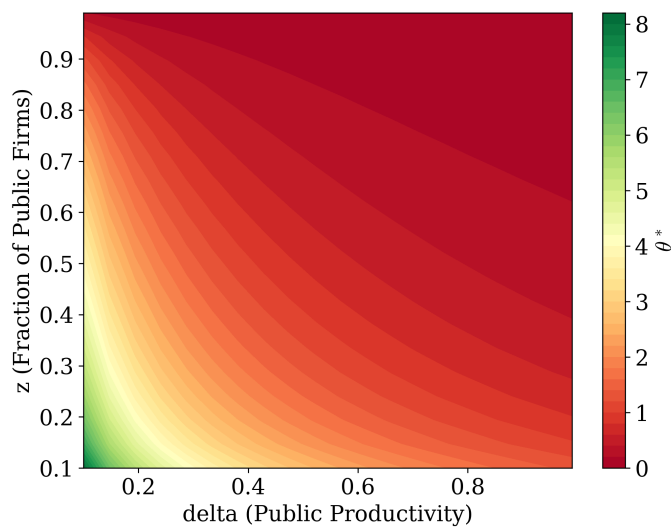


Figure 9. Parameter space This figure describes equilibriums that exist when varying in the parameter space. Figures vary δ and z parameters. Assumes $\lambda = 1/2$, $\ln \theta_i \sim N(\mu = 0, \sigma = 1)$, and $R(\theta^*) = \rho * E[\theta_i | \theta_i > \theta^*]$ where scaling factoring $\rho = 0.1$. Panel (a) presents the regions in the parameter space where a separating equilibrium will exist. Panel (b) shows how the cutoff type θ^* changes in the parameter space.

(a) Regions where a separating equilibrium exists



(b) θ^* as a function of δ and z



9 Tables

Table 1. Summary statistics

This table presents pooled summary statistics from the main difference-in-differences specification. The unit of observation is firm-quarters from 2020 through 2024. The sample includes GitHub organizations with successful matches to PitchBook VC transaction history. $\% \Delta$ Employees is the percentage change in number of employees. *Code Changes* is the cumulative number of lines of code changed in the project, measured as the sum of code *additions* and *deletions*. *External Share* is the share of code changes by developers with status of “no status” or contributor, divided by total code changes. *# Unique Contributors* is the number of unique contributors over the lifetime of a project. $1(\text{Crypto})$ is an indicator variable for firms in the crypto sector, as determined through PitchBook keywords. $1(\text{Treat})$ is an indicator for firms with prior ties to FTX-affiliated investors. $1(\text{Post})$ is an indicator for periods after June 2022.

	count	mean	sd	min	25%	50%	75%	max
Num People	123,605	401.50	3,593.07	1.	10	31	110	374,638
$\% \Delta$ Employees	123,605	0.04	0.23	-0.90	-0.00	0.00	0.06	15.00
Code Changes	123,605	159,881	1,154,993.19	1	366	5,223	45,521	67M
% External Share	123,605	0.51	0.44	0.00	0.01	0.51	1.00	1.00
# Unique Contrib	123,605	13.25	59.77	1	2	4	10	8,198
$1(\text{Crypto})$	123,605	0.12	0.32	0	0	0	0	1
$1(\text{Post})$	123,605	0.57	0.49	0	0	1	1	1
$1(\text{Treat})$	123,605	0.12	0.33	0	0	0	0	1

Table 2. Relationship between open-source and salaries

This table presents estimates on the effect of prior work history on future compensation. The unit of observation is at the developer-position level. *External share_{i,p}* aggregates the share of work by developer *i* that was made as contributions to a repository where they were not a formal collaborator during the time of employment at position *p*, where *p* indexes positions for developer *i*. *log(Next Comp)* is the compensation at their next position (i.e., *p* + 1). Column 1 includes fixed effects for the firm of position *p*, as well as the state and seniority level for position *p* as provided by Revelio. This also includes date fixed effects, or the month in which the worker left position *p*. Column 2 uses substitutes Firm for Developer fixed effects. Controls include log(Compensation), duration of employment, and volume of work (measured as log number of pull request changes) at position *p*. Standard errors in parentheses are clustered by developer * *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01.

$$\log(\text{Compensation of Next Position})_{i,p+1} = \alpha + \beta \text{External share}_{i,p} + \Gamma_{i,p} + \varepsilon_{i,p}$$

	(1) log(Next Comp)	(2) log(Next Comp)
External Share	0.0190*** (0.00569)	0.0192*** (0.00725)
Dev FE		✓
Firm FE	✓	
Month FE	✓	✓
State FE	✓	✓
Seniority FE	✓	✓
Controls	✓	✓
Observations	87580	95043
R-squared	0.315	0.426
Number of Clusters	47441	30848

Table 3. Hazard model: next round of fundraising

This table reports hazard ratios from Cox proportional-hazards models that estimate how quickly startups secure their next funding round. The variable *post* is an indicator encoded as 1 after the November 2022 FTX collapse. *Treat* is an indicator for firms with prior ties to FTX-affiliated investors. Robust standard errors clustered at the firm level are displayed in parentheses. Column (2-4) include fixed effects. Column (4) restricts the sample to only crypto firms. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$h_{is}(t) = h_0(t) \exp \left(\beta_1 1(\text{Post})_{is} + \beta_2 1(\text{Treat})_i + \beta_3 (1(\text{Post})_{is} \times 1(\text{Treat})_i) \right. \\ \left. + \sum_r \gamma_r \mathbf{1}\{\text{VC Round} = r\} + \sum_j \delta_j \mathbf{1}\{\text{Industry} = j\} \right)$$

	(1)	(2)	(3)	(4)
1(Post)	-0.133*** (0.0457)	-0.156*** (0.0507)	-0.140*** (0.0508)	-0.299** (0.137)
1(Treat)	0.610*** (0.0678)	0.696*** (0.0769)	0.865*** (0.0906)	0.769*** (0.109)
1(Post) $\times$ 1(Treat)	-0.424*** (0.127)	-0.541*** (0.133)	-0.551*** (0.133)	-0.298* (0.175)
Sample exp(β_3) - 1	All -34.6%	All -41.8%	All -42.4%	Crypto-only -25.8%
VC-round FE		✓	✓	✓
Industry FE		✓	✓	✓
Crypto $\times$ Post FE		✓	✓	✓
Obs	9477	8200	8200	1818
Wald χ^2	102.2	2789.9	2777.0	3384891.7
Log-lik	-28250.6	-22843.3	-22835.9	-3983.1

Table 4. Difference-in-differences: labor substitution

This table presents estimates of the effect of the FTX collapse on employment growth and share of work complete by external developers. The sample includes firms with GitHub repositories from 2022 through June 2024. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post FTX* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variables are: (1-2) *Employee Growth* - the percentage change in employment; and (3-4) *External Share Volume* - the share of pull requests (by volume, or sum of code additions and deletions) contributed by external contributors; this model is weighted by lines of code such that the coefficient based on the average across lines of code instead of firms. All specifications include firm and date fixed effects, with standard errors clustered at the firm and date level. Controls in column 2 and 4 include age (absorbed in two-way fixed effects), age squared, and size. Column 2 and 4 also have 1(Post) interactions with fixed effects for crypto-sector, industry, and geography defined at the state-level in the U.S. and country-level elsewhere. Crypto fixed effects are separate from industry fixed effects, and the dummy is determined based on PitchBook keyword designations. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{Treat})_i) + \chi_i + \delta_t + \varepsilon_{i,t}$$

	%Δ Employees		% External Share	
	(1)	(2)	(3)	(4)
1(Post) × 1(Treat)	-0.0618*** (0.00558)	-0.0237*** (0.00538)	0.105*** (0.0316)	0.102*** (0.0320)
Firm FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Post × Industry FE		✓		✓
Post × Crypto FE		✓		✓
Post × Location FE		✓		✓
Controls		✓		✓
Observations	122174	111282	122174	111282
Adj R-squared	0.110	0.0994	0.640	0.650
Number of Clusters	11718	10496	11718	10496

Table 5. Difference-in-differences: repository activity

This table presents estimates of the effect of the FTX collapse on project statistics. The outcome y variables are: (1) log cumulative number of unique contributors in each quarter, (2) log cumulative number of public projects by the firm, (3) log number of lines of code changed in each quarter, measured as the sum of additions and deletions, and (4) log cumulative number of stars. $Post$ is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. Standard errors in parentheses are clustered at the firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta_1(1(Post)_t \times 1(Treat)_i) + \chi_i + \delta_t + \varepsilon_{i,t}$$

	(1) log(Contributors)	(2) log(Projects)	(3) log(Code Volume)	(4) log(Stars)
post=1 $\times$ treat=1	0.133*** (0.0237)	0.152*** (0.0291)	0.166*** (0.0580)	0.172*** (0.0391)
Firm FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Other Post $\times$ FE	✓	✓	✓	✓
Observations	98896	121123	121123	121123
R-squared	0.927	0.780	0.864	0.962
Number of Clusters	9358	11629	11629	11629

Table 6. Difference-in-differences: developer composition

This table presents estimates of the effect of the FTX collapse on the characteristics of developers that self-select to contributing towards impacted organizations. The sample includes firms with GitHub repositories from 2022 through June 2024. The unit of observation is at the firm-month-developer level. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. *Developer Merge Rate* is the acceptance rate of a developer making a contribution. *Has Top Contributions* is an indicator variable for whether the developer has had a contribution accepted in a top 100 organization. Standard errors in parentheses are clustered by firm * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{Treat})_i) + \chi_i + \delta_t + \varepsilon_{i,t}$$

	(1) Developer Merge Rate	(2) Has Top Contributions
1(Post) FTX $\times$ 1(Treat)	-0.013*** (0.004)	-0.133*** (0.023)
Firm FE	✓	✓
Date FE	✓	✓
Estimator	OLS	Binomial
N	2,044,113	2,044,113
R^2 (OLS) / Pseudo- R^2 (Binomial)	0.168	0.810

Table 7. Signaling effectiveness for promotion

This table presents estimates of the effect of the FTX collapse on the likelihood external contributions are associated with being “hired,” or being promoted from external to internal. $\log(\text{volume})(\text{external})_{i,t}$ are contributions made by worker i as an external contributor in month t . The dependent variable $1(\text{Promoted})_{i,t \rightarrow t+6}$ indicates if the worker makes a contribution as an internal member to any project in the next six months, multiplied by 100 for scaling. The specification includes worker and date fixed effects. Controls include log of total contributions and acceptance rates. Standard errors in parentheses are clustered by worker * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$1(\text{Promoted})_{i,t \rightarrow t+6} = \beta_0 + \beta_1 \log(\text{volume}_{\text{external}})_{i,t} + \beta_2 \log(\text{volume}_{\text{total}})_{i,t} + \beta_3 (\text{acceptance rate})_{i,t}$$

	(1) $1(\text{Promoted})_{\times 100}$	(2) $1(\text{Promoted})_{\times 100}$ (Same Firm)	(3) $1(\text{Promoted})_{\times 100}$ (Diff Firm)
log(External Code Changes)	0.427*** (0.021)	0.393*** (0.021)	0.034*** (0.004)
log(Total Code Changes)	0.061*** (0.018)	0.208*** (0.018)	-0.147*** (0.005)
Merge Rate	-1.215*** (0.064)	-1.129*** (0.062)	-0.086*** (0.022)
Worker FE	Yes	Yes	Yes
Date FE	Yes	Yes	Yes
N	13,165,304	13,165,304	13,165,304
Adjusted R^2	0.182	0.139	0.048
Dep. Mean	7.418	5.461	1.956

Table 8. Signaling effectiveness for treated firms

This table presents estimates of the effect of the FTX collapse on the likelihood external contributions are associated with being “hired,” or being promoted from external to internal. $\log(\text{volume})(\text{external}, \text{treated})_{i,t}$ are contributions made by worker i to treated firms made as an external contributor in month t . Other variables are analogously constructed. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. $1(\text{Post})$ is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variable $1(\text{Promoted})_{i,t \rightarrow t+6}$ indicates if the worker makes a contribution as an internal member to any project in the next six months, multiplied by 100 for scaling. The specification includes worker and date fixed effects. Controls include log of total contributions and merge rate terms and their interaction with $1(\text{Post})$. Standard errors in parentheses are clustered by worker * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$\begin{aligned}
1(\text{Promoted})_{i,t \rightarrow t+6} = & \beta_0 + \beta_1 \log(\text{volume}_{\text{external}, \text{treated}})_{i,t} + \beta_2 \log(\text{volume}_{\text{external}, \text{control}})_{i,t} \\
& + \beta_3 (\text{post}_t \times \log(\text{volume}_{\text{external}, \text{treated}})_{i,t}) \\
& + \beta_3 (\text{post}_t \times \log(\text{volume}_{\text{external}, \text{control}})_{i,t}) + \chi_i + \delta_t + \varepsilon_{i,t}
\end{aligned}$$

	(1)	(2)
	1(Promoted) $\times 100$	
log(External Treated)	0.707*** (0.038)	0.705*** (0.048)
log(External Control)	0.413*** (0.020)	0.255*** (0.035)
1(Post) $\times$ log(External Treated)		0.012 (0.067)
1(Post) $\times$ log(External Control)		0.343*** (0.044)
Worker FE	✓	✓
Date FE	✓	✓
Controls	✓	✓
N	13,165,304	13,165,304
R^2	0.286	0.287

Table 9. Signaling effectiveness in differences and split by hiring

This table presents estimates of the effect of the FTX collapse on the likelihood external contributions are associated with being “hired,” or being promoted from external to internal. $\log(\text{volume})(\text{external}, \text{treated})_{i,t}$ are contributions made worker i to treated firms made as an external contributor in month t . Other variables are analogously constructed. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. $1(\text{Post})$ is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variable $1(\text{Promoted})_{i,t \rightarrow t+6}$ indicates if the worker makes a contribution as an internal member to any project in the next six months, multiplied by 100 for scaling. The independent variable Δ is the difference between volume to treated firms and control firms. Column 2 isolates hiring to a firm that a worker made contributions and column 3 isolates hiring to any other firm. The specification includes worker and date fixed effects. Controls include log of total contributions and merge rate terms and their interaction with $1(\text{Post})$. Standard errors in parentheses are clustered by worker * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$\Delta_{it} = \log(\text{volume}_{\text{external}, \text{treated}})_{i,t} - \log(\text{volume}_{\text{external}, \text{control}})_{i,t}$$

$$1(\text{Promoted})_{i,t \rightarrow t+6} = \beta_0 + \beta_1 \Delta_{it} + \beta_2 (\text{post}_t \times \Delta_{it}) + \chi_i + \delta_t + \varepsilon_{i,t}$$

	(1) 1(Promoted) $\times 100$	(2) 1(Promoted) $\times 100$ (Same Firm)	(3) 1(Promoted) $\times 100$ (Diff Firm)
Δ	-0.158*** (0.029)	-0.150*** (0.027)	-0.008 (0.006)
$1(\text{Post}) \times \Delta$	-0.291*** (0.033)	-0.235*** (0.033)	-0.056*** (0.006)
Worker FE	✓	✓	✓
Date FE	✓	✓	✓
Controls	✓	✓	✓
N	13,165,304	13,165,304	13,165,304
R^2	0.286	0.249	0.170
Dep. Mean	7.418	5.461	1.956

Table 10. Difference-in-differences: firm outcomes

This table presents estimates of the effect of the FTX collapse on firm outcomes after the FTX shock. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post FTX* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. Standard errors in parentheses are clustered by firm. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{Treat})_i) + \beta_2 1(\text{Treat})_i + \varepsilon_{i,t}$$

(a) Corporate Outcome					
	(1) Bankruptcy $\times 100$	(2) IPO $\times 100$	(3) M&A $\times 100$	(4) Secondary $\times 100$	(5) VC $\times 100$
1(Post) $\times$ 1(Treat)	-0.0152 (0.0156)	-0.0668*** (0.0154)	0.0556* (0.0291)	0.00749 (0.0491)	-2.832*** (0.127)
Date FE	✓	✓	✓	✓	✓
Mean of DV	0.0773	0.0246	0.229	0.231	2.357
Observations	605466	605466	605466	605466	605466
R-squared	0.000253	0.000356	0.000169	0.00105	0.00489
Number of Clusters	11448	11448	11448	11448	11448

(b) Reported Outcome from VC Funding Round			
	(1) Up Round $\times 100$	(2) Flat Round $\times 100$	(3) Down Round $\times 100$
1(Post) $\times$ 1(Treat)	-2.422*** (0.161)	-0.00187 (0.0643)	0.0388 (0.0412)
Date FE	✓	✓	✓
Mean of DV	1.642	0.329	0.133
Observations	153909	153909	153909
R-squared	0.00706	0.000125	0.000165
Number of Clusters	2914	2914	2914

A Appendix: Figures

Figure IA.1. Comparison of internal vs. external developers

This plot shows the histogram of sophistication ratings of developer contributions.

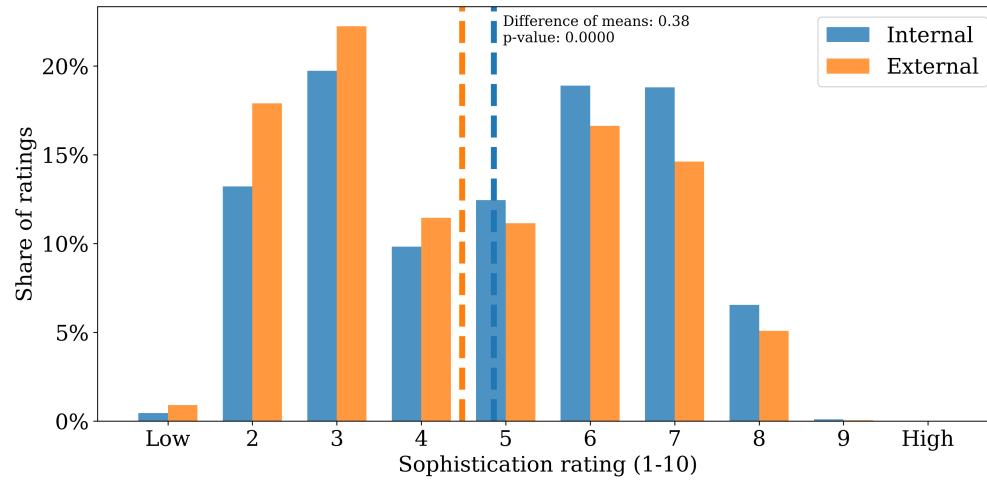


Figure IA.2. VC impact on share of open-soiurce development

This figure shows the relationship between VC investment and share of open-source code. The bars plots the t -statistics on γ_k from the following regression. The panel regression is from 2018-2021 and the unit of observation is a firm-quarter. The indicator variable encodes 1 when investor k has participated in a venture capital funding round. These indicator variables are not mutually exclusive and therefore these are not “fixed effects” in that there is no “base” investor excluded. These include the top 50 investors in the sample by number of deals. The dashed lines indicate a t -statistic of more than 2.5 or less than -2.5.

$$\log(\text{External Code}_{it}) = \beta_1 \log(\text{Total Code}_{it}) + \sum_k \gamma_k 1(\text{Has Investor } k)_{k,it} + \alpha_i + \delta_t + \varepsilon_{i,t}$$

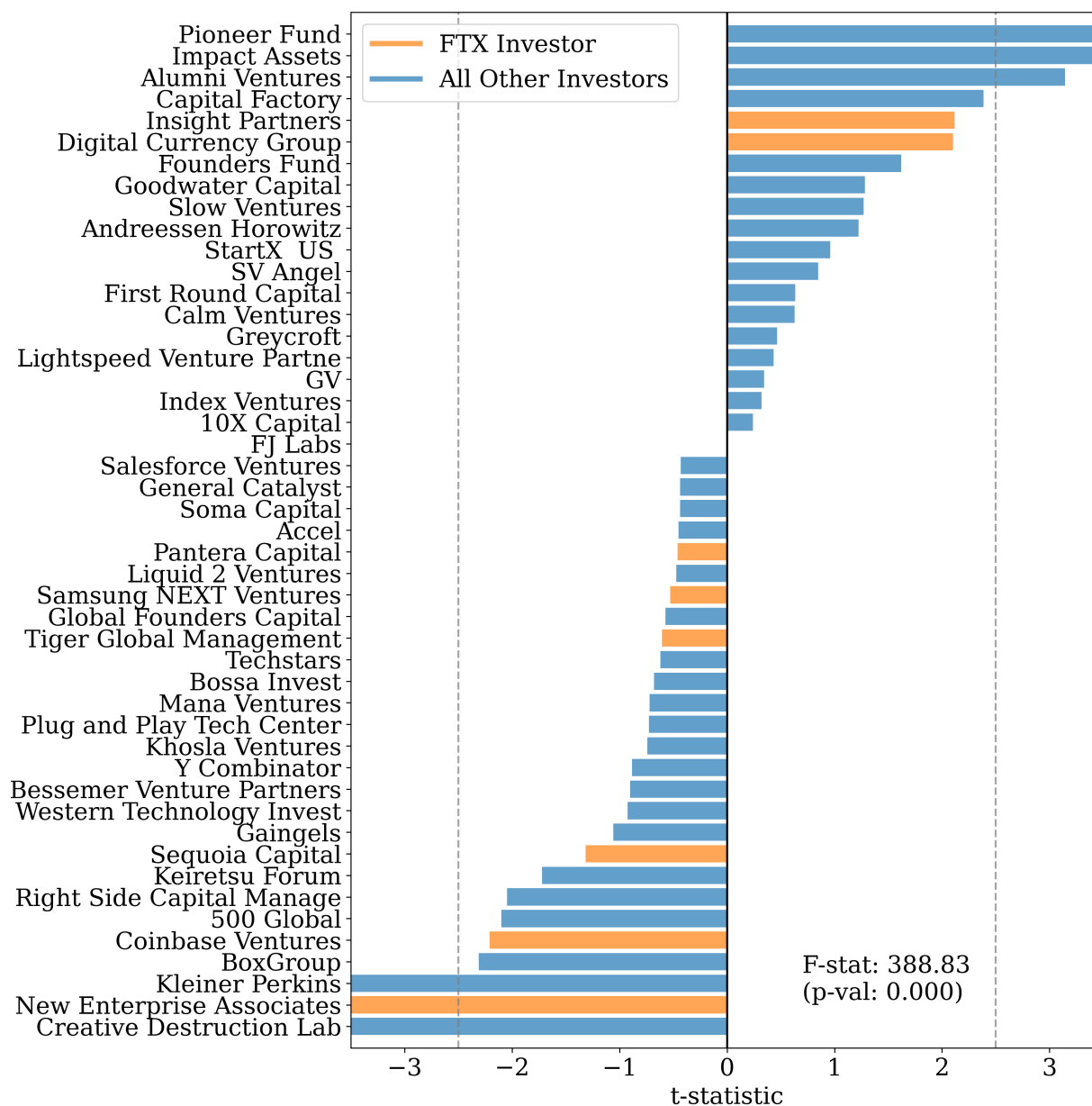


Figure IA.3. Histogram: Primary Language, as a share of whole repository

This figure studies the primary language in each repository. It plots a histogram of the share of each repo that the primary language occupies.

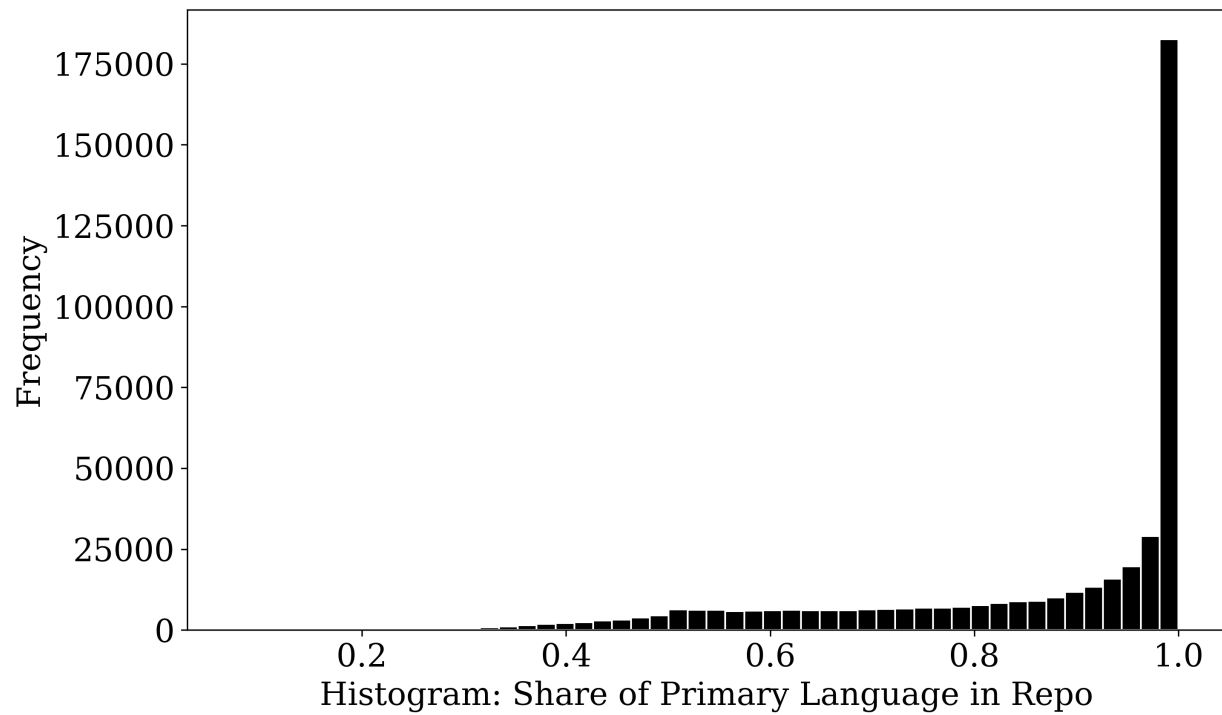


Figure IA.4. Test statistics by programming language

This figure presents average size of the treated and control groups and their average growth rate

$$y_{i,t} = \beta_0 + \sum_l \beta_l \text{post}_t \times \text{treat}_i \times (\text{Primary Language} == l)_i + \chi_i + \delta_t + \varepsilon_{i,t}$$

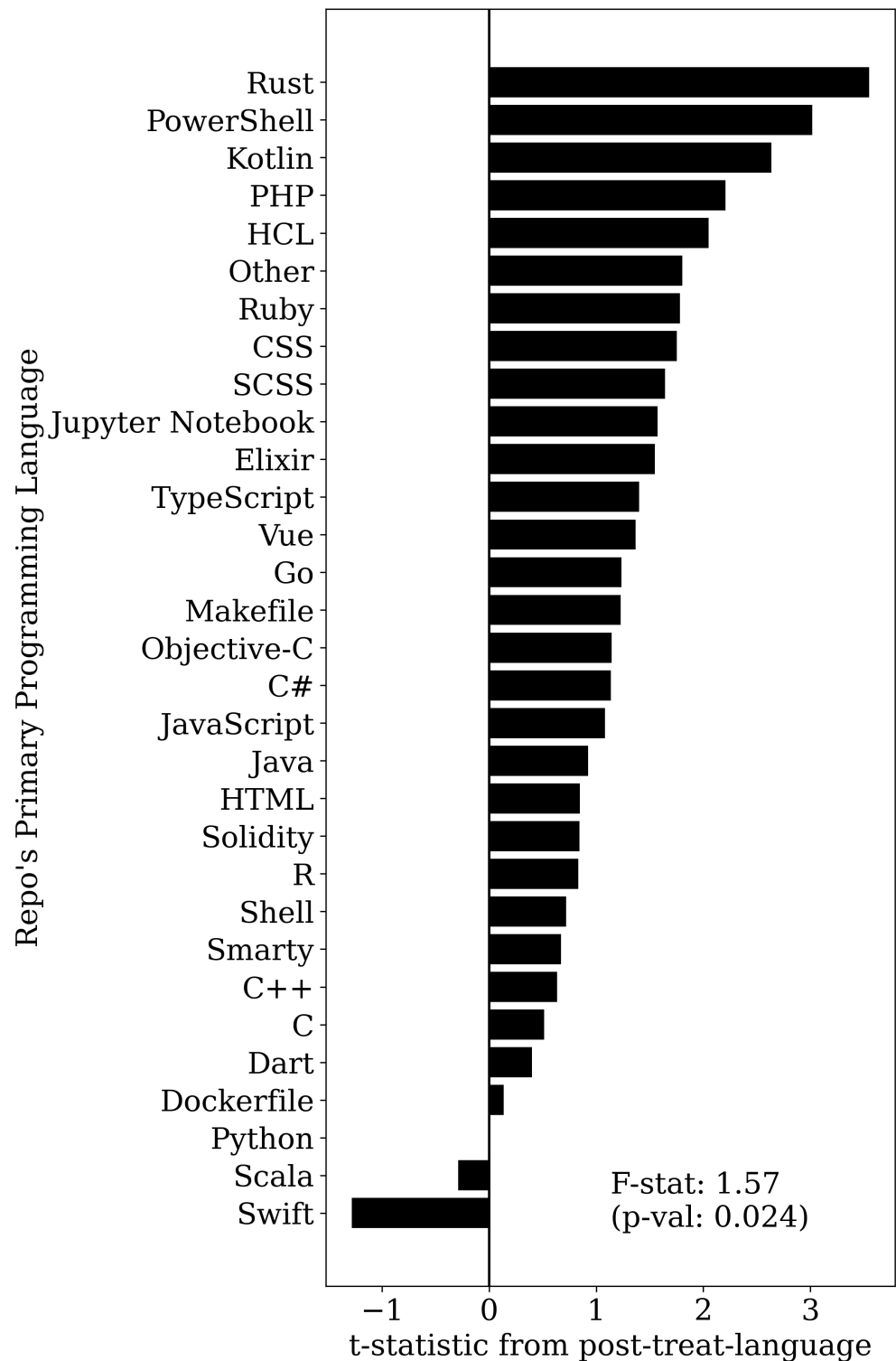
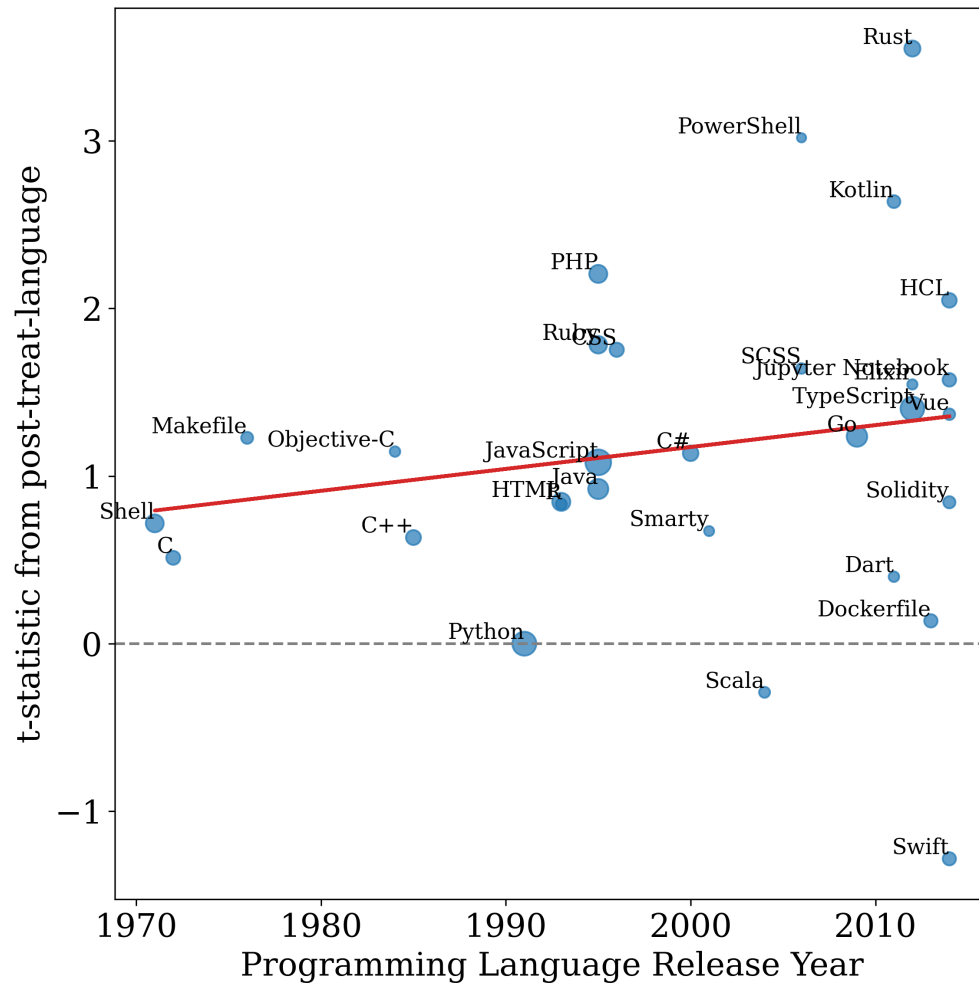


Figure IA.5. Test statistics by programming language

This figure presents average size of the treated and control groups and their average growth rate

$$y_{i,t} = \beta_0 + \sum_l \beta_l \text{post}_t \times \text{treat}_i \times (\text{Primary Language} == l)_i + \chi_i + \delta_t + \varepsilon_{i,t}$$



B Appendix: Tables

Table IA.I. Author summary

This table presents a summary of author types. Total PRs is total pull requests, or the number of pull request contributions. Total lines changed is the sum of additions and deletions in the code differences between versions. Column 3 divides Column 2 by Column 1. Column 4 shows the number of contributions that are accepted by the project administrators and Column 5 shows the same idea as a share of lines of code changed, as in Column 2.

	(1)	(2)	(3)	(4)	(5)
Author Association	Total PRs	Total Lines Changed	Avg Lines Changed	Pct PRs Merged	Pct Lines Merged
Owner	83,608	307,264,389	3,675.1	84.55%	78.36%
Member	6,346,951	16,767,292,902	2,641.8	82.15%	68.39%
Collaborator	5,479,209	22,974,937,063	4,193.1	81.80%	68.09%
Contributor	12,698,590	30,448,538,873	2,397.8	77.69%	60.84%
None	4,289,008	22,904,484,263	5,340.3	50.47%	29.06%

Table IA.II. Relationship between tenure, prior education, and external contributions

This table describes the relationship between probability a contribution is made as an external developer and years of experience. The unit of observation is one contributions c , or a pull request, by developer i in month t . The y variable is an indicator variable encoded as 1 if the work was completed as an external contributor. The regression includes weight by total changes in lines of code. Years of Experience (YoE) is colinear when including worker and date fixed effects and therefore do not have coefficients. Standard errors in parentheses are clustered by developer and date * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$1(\text{External})_{i,c,t} = \beta_0 + \beta_1(1(\text{Not Top School})_i \times \text{Years of Experience}_{i,t}) + \chi_i + \delta_t + \varepsilon_{i,c,t}$$

	1(External)
1(Not Top School) $\times$ YoE	-0.019*** (0.007)
Worker FE	Yes
Date FE	Yes
N	3,543,455
Adjusted R^2	0.801

Table IA.III. Firm outcomes: employee turnover

This table presents the relationship between external share and a firm's employment dynamics. The dependent variables are hiring rate (number of new employees / total employees), separating rate (number of employees who changed employment status / total employees), and headcount growth (percentage change in employees). External share is the contemporaneous amount of work in a firm that is completed by developers who are not formal collaborators. All columns include firm and date fixed effects. Standard errors in parentheses are clustered by firm * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \alpha + \beta \text{External share}_{i,t} + \chi_i + \delta_t + \varepsilon_{i,t}$$

	(1) % Hiring Rate	(2) % Separation Rate	(3) %Δ Employees
% External Share	0.00210*** (0.000531)	-0.000316 (0.000502)	0.00259*** (0.000819)
Firm FE	✓	✓	✓
Month FE	✓	✓	✓
Observations	146154	146154	146154
R-squared	0.389	0.281	0.138
Number of Clusters	6838	6838	6838

Table IA.IV. Difference-in-differences: labor substitution by days since prior funding

This table presents estimates of the effect of the FTX collapse on employment growth and share of work complete by external developers. The sample includes firms with GitHub repositories from 2022 through June 2024. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post FTX* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variables are: (1-2) *Employee Growth* - the percentage change in employment; and (3-4) *External Share Volume* - the share of pull requests (by volume, or sum of code additions and deletions) contributed by external contributors; this model is weighted by lines of code such that the coefficient based on the average across lines of code instead of firms. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta(1(\text{Post})_t \times 1(\text{Treat})_i \times 1(\text{Days Since Prior Funding} > 500)_i) + \chi_i + \delta_t + \varepsilon_{i,t}$$

	(1) %Δ Employees	(2) % External Share
1(Post) × 1(Treat)	-0.0169*** (0.00501)	0.0701* (0.0390)
1(Post) × 1(Days Since Prior Funding > 500)	0.0531*** (0.00304)	-0.0235 (0.0321)
1(Post) × 1(Treat) × 1(Days Since Prior Funding > 500)	-0.0419*** (0.00688)	0.122* (0.0727)
Firm FE	✓	✓
Date FE	✓	✓
Other FE	✓	✓
Observations	122174	122174
R-squared	0.150	0.641
Number of Clusters	11718	11718

Table IA.V. Difference-in-differences: labor substitution heterogeneity

This table presents estimates of the effect of the FTX collapse on employment growth and share of work complete by external developers. The sample includes firms with GitHub repositories from 2022 through June 2024. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post FTX* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variable *External Share* is the share of pull requests (by volume, or sum of code additions and deletions) contributed by external contributors. This model is weighted by lines of code such that the coefficient based on the average across lines of code instead of firms. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta(1(\text{Post})_t \times 1(\text{Treat})_i \times \% \Delta \text{Employees}_{i,t}) + \beta \% \Delta \text{Employees}_{i,t} + \chi_i + \delta_t + \varepsilon_{i,t}$$

	(1) % External Share
1(Treat) $\times$ 1(Post)	0.104*** (0.0339)
% Δ Employees	-0.0335 (0.0533)
1(Treat) $\times$ % Δ Employees	-0.0300 (0.108)
1(Post) $\times$ % Δ Employees	0.137** (0.0685)
1(Treat) $\times$ 1(Post) $\times$ % Δ Employees	-0.382** (0.152)
Firm FE	✓
Date FE	✓
Other FE	✓
Observations	122174
R-squared	0.641
Number of Clusters	11718

Table IA.VI. Difference-in-differences: labor substitution

This table presents estimates of the effect of the FTX collapse on employment growth and share of work complete by external developers. The sample includes firms with GitHub repositories from 2022 through June 2024. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variable $\log(\text{Code from External})$ is the log of volume of pull requests (by volume, or sum of code additions and deletions) contributed by external contributors. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$\log(\text{Code from External})_{i,t} = \beta_0 + \beta_1 \log(\text{Total Code})_{i,t} + \chi_i + \delta_t + \varepsilon_{i,t}$$

	(1)	(2)
	$\log(\text{Code from External})$	
$1(\text{Post}) \times 1(\text{Treat})$	0.334*** (0.0801)	0.161*** (0.0577)
$\log(\text{Total Code})$		0.695*** (0.00485)
Firm FE	✓	✓
Date FE	✓	✓
Observations	132951	132951
R-squared	0.539	0.685
Number of Clusters	13010	13010

Table IA.VII. Difference-in-differences: code sophistication

This table presents estimates of the effect of the FTX collapse on code sophistication. The sample includes firms with GitHub repositories from 2022 through June 2024. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variable *Rating* is the average sophistication of code by volunteers for organisation i at time t , weighted by the lines of code. Sophistication is rated as an integer from 0-100 by an OpenAI Chat Completion assistant. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$\text{Rating}_{i,t} = \beta_0 + \beta_1(\text{post}_t \times \text{treat}_i) + \chi_i + \gamma_j + \delta_t + \varepsilon_{i,t}$$

	(1) Rating	(2) Rating
1(Treat) $\times$ 1(Post)	4.627*** (1.740)	4.869*** (1.843)
Firm FE	✓	✓
Date FE	✓	✓
Controls		✓
Observations	39403	36829
R-squared	0.267	0.268
Number of Clusters	5640	5191

Table IA.VIII. Signaling effectiveness

This table presents estimates of the effect of the FTX collapse on the likelihood external contributions are associated with being “hired,” or being promoted from external to internal. This is analogous to Table 8 except the independent variables are now: contributions as a share of all monthly contributions. $\% \text{ share}(\text{external}, \text{treated})_{i,t}$ are contributions made worker i to treated firms made as an external contributor, as a share of total work in month t . Other shares are similarly constructed. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. $1(\text{Post})$ is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variable $1(\text{Hired})_{i,t \rightarrow t+3}$ indicates if the worker makes a contribution as an internal member to any project in the next three months, multiplied by 100 for scaling. The specification includes worker and date fixed effects. Standard errors in parentheses are clustered by worker * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$\begin{aligned}
1(\text{Promoted})_{i,t \rightarrow t+6} &= \beta_0 + \beta_1 \% \text{ share}(\text{external}, \text{treated})_{i,t} + \beta_2 \% \text{ share}(\text{external}, \text{control})_{i,t} \\
&\quad + \beta_3 (\text{post}_t \times \% \text{ share}(\text{external}, \text{treated})_{i,t}) \\
&\quad + \beta_3 (\text{post}_t \times \% \text{ share}(\text{external}, \text{control})_{i,t}) + \chi_i + \delta_t + \varepsilon_{i,t} \\
\Delta_{it} &= \% \text{ share}(\text{external}, \text{treated})_{i,t} - \% \text{ share}(\text{external}, \text{control})_{i,t} \\
1(\text{Promoted})_{i,t \rightarrow t+6} &= \beta_0 + \beta_1 \Delta_{it} + \beta_2 (\text{post}_t \times \Delta_{it}) + \chi_i + \delta_t + \varepsilon_{i,t}
\end{aligned}$$

	(1)	(2)	(3)	(4)
	1(Promoted)	1(Promoted) (Same Company)	1(Promoted) (Different Company)	
$\%(\text{ext}, \text{treat})$	0.113*** (0.040)	0.052 (0.047)	0.182*** (0.039)	
$\%(\text{ext}, \text{control})$	-0.144*** (0.032)	-0.303*** (0.037)	0.246*** (0.017)	
$\%(\text{ext}, \text{treat}) \times 1(\text{Post})$	0.215*** (0.065)	0.285*** (0.079)	0.072 (0.061)	
$\%(\text{ext}, \text{control}) \times 1(\text{Post})$	0.366*** (0.052)	0.447*** (0.061)	0.110*** (0.026)	
Δ				-0.202*** (0.015)
$\Delta \times 1(\text{Post})$				-0.091*** (0.023)
Worker FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Dep. Mean	0.074	0.055	0.02	0.02
Marginal Effects: β_3	0.0534	0.0705	0.0178	
Marginal Effects: $\Delta \times 1(\text{Post})$				-0.0224
N	13,165,304	13,165,304	13,165,304	13,165,304
R^2	0.419	0.402	0.437	0.437

Table IA.IX. Signaling effectiveness, by ratings

This table presents estimates of the effect of the FTX collapse on the likelihood external contributions are associated with being “hired,” or being promoted from external to internal. This is analogous to Table 8 with the addition of code ratings. $\% \text{ share}(\text{external}, \text{treated})_{i,t}$ are contributions made worker i to treated firms made as an external contributor, as a share of total work in month t . Other shares are similarly constructed. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. $1(\text{Post})$ is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variable $1(\text{Hired})_{i,t \rightarrow t+3}$ indicates if the worker makes a contribution as an internal member to any project in the next three months, multiplied by 100 for scaling. The specification includes worker and date fixed effects. Standard errors in parentheses are clustered by worker * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$\begin{aligned}
1(\text{Promoted})_{i,t \rightarrow t+6} = & \beta_0 + \beta_1 \log(\text{volume}_{\text{external}, \text{treated}})_{i,t} + \beta_2 \log(\text{volume}_{\text{external}, \text{control}})_{i,t} \\
& + \beta_3 (\text{post}_t \times \log(\text{volume}_{\text{external}, \text{treated}})_{i,t}) \\
& + \beta_3 (\text{post}_t \times \log(\text{volume}_{\text{external}, \text{control}})_{i,t}) + \chi_i + \delta_t + \varepsilon_{i,t}
\end{aligned}$$

	1(Promoted) $\times 100$	
	(1)	(2)
Δ	0.687*** (0.052)	0.444*** (0.052)
Avg Rating	0.019*** (0.006)	0.012** (0.006)
$1(\text{Post}) \times \Delta$	-3.280*** (0.061)	-3.238*** (0.061)
$\log(\text{Total Code Changes})$		0.307*** (0.005)
Worker FE	Yes	Yes
Date FE	Yes	Yes
N	13,165,304	13,165,304
Adjusted R^2	0.181	0.181
Dep. Mean	7.418	7.418

Table IA.X. Difference-in-differences: characteristics of new workers

This table presents estimates of the effect of the FTX collapse on the characteristics of developers that start at the firm after the FTX shock. The sample includes all developer work histories. The unit of observation is at the developer-position level. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The estimate includes fixed effects for if the position is at firm j , has start date t , and for developer i . Standard errors in parentheses are clustered by firm. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,p} = \beta_0 + \beta_1(\text{post}_t \times \text{treat}_i) + \chi_i + \gamma_j + \delta_t + \varepsilon_{i,p}$$

(a) Characteristics of new hires: wage and level

	(1) log(Comp)	(2) Seniority	(3) Position Number
1(Post) $\times$ 1(Treat)	-0.023*** (0.002)	-0.006* (0.003)	0.008 (0.089)
Firm FE	✓	✓	✓
Worker FE	✓	✓	✓
Start Date FE	✓	✓	✓
N	6,656,271	6,656,344	6,656,344
R^2	0.812	0.672	0.854

(b) Characteristics of new hires: location

	(1) 1(US)	(2) 1(California)	(3) 1(Top 5 Metro)	(4) 1(Modal State)
1(Post) $\times$ 1(Treat)	-0.032*** (0.001)	-0.090*** (0.001)	-0.057*** (0.001)	-0.092*** (0.001)
Firm FE	✓	✓	✓	Yes
Date FE	✓	✓	✓	✓
N	8,037,553	7,366,992	7,874,836	7,363,594
R^2	0.396	0.314	0.328	0.208

Table IA.XI. Difference-in-differences: characteristics of future jobs of existing workers

This table presents estimates of the effect of the FTX collapse on the characteristics of developers that start at the firm after the FTX shock. The sample includes all developer work histories. The unit of observation is at the developer-position level. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The estimate includes fixed effects for if the position is at firm j , has start date t , and for developer i . Standard errors in parentheses are clustered by worker. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,p} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{At Treated Firm Prior to 2022})_i) + \chi_i + \delta_t + \varepsilon_{i,p}$$

	<u>log(Comp)</u>	<u>1(Founder)$\times 100$</u>	<u>1(Open Source)$\times 100$</u>
	(1)	(2)	(3)
1(Post) $\times$ 1(At Treated Firm Prior to 2022)	-0.029*** (0.002)	0.086*** (0.019)	-0.215*** (0.034)
Worker FE	✓	✓	✓
Start Date FE	✓	✓	✓
Dep. Mean	11.081	0.134	0.687
N	6,657,755	5,639,188	5,639,188
R^2	0.799	0.436	0.349

Table IA.XII. Difference-in-differences: labor substitution heterogeneity

This table presents estimates of the effect of the FTX collapse on employment growth and share of work complete by external developers. The sample includes firms with GitHub repositories from 2022 through June 2024. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post FTX* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variable *External Share* is the share of pull requests (by volume, or sum of code additions and deletions) contributed by external contributors. This model is weighted by lines of code such that the coefficient based on the average across lines of code instead of firms. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)
	% Δ Employees	% External Share
1(Post) $\times$ 1(Treat)	-0.0250*** (0.00563)	0.121*** (0.0329)
Firm FE	✓	✓
Date FE	✓	✓
Post $\times$ Crypto FE	✓	✓
Post $\times$ Industry FE	✓	✓
Post $\times$ Location FE	✓	✓
VC Indicators	✓	✓
Controls	✓	✓
Observations	111282	111282
R-squared	0.0995	0.652
Number of Clusters	10496	10496

Table IA.XIII. Difference-in-differences: matching design

This table presents estimates of the effect of the FTX collapse on employment growth and share of work complete by external developers. Treated firms are matched to control firms with replacement. Firms must be within the same primary industry group, within the same state if in the U.S. and within the same country if outside of the U.S., be within a 5 year difference in age, and within 25% in number of contributors as of 2021. The specifications include cohort, firm, and date fixed effects. Standard errors in parentheses are clustered by cohort * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,c,t} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{Treat})_i) + \chi_i + \delta_t + \eta_c + \varepsilon_{i,c,t}$$

	(1)	(2)
	Empl Growth	External Share Volume
1(Post) $\times$ 1(Treat)	-0.0404*** (0.00531)	0.0859*** (0.0271)
Cohort FE	✓	✓
Firm FE	✓	✓
Date FE	✓	✓
Observations	329178	329178
R-squared	0.164	0.661
Number of Clusters	1235	1235

Table IA.XIV. Difference-in-differences: heterogeneity

This table presents estimates of the effect of the FTX collapse on employment growth and share of work complete by external developers. The sample includes firms with GitHub repositories from 2022 through June 2024. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post FTX* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variables are: (1) *Employee Growth* - the percentage change in employment; and (2) *External Share Volume* - the share of pull requests (by volume, or sum of code additions and deletions) contributed by external contributors. All specifications include firm and date fixed effects, with standard errors clustered at the firm and date level. Indicators for firm characteristics are formed as of December 31, 2021. *Small firms* are firms with less than the treated group median (167 employees) as of December 31, 2021. *Young firms* are firms that are less than the treated group median (five years) as of the date of December 31, 2021. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1) % Δ Employees	(2) Δ External Share
1(Post) $\times$ 1(Treat)	-0.0275*** (0.00413)	0.142*** (0.0543)
1(Post) $\times$ 1(Young Firm)	-0.0770*** (0.00480)	0.0365 (0.0287)
1(Post) $\times$ 1(Treat) $\times$ 1(Young Firm)	0.0237** (0.0111)	0.0573 (0.0777)
1(Post) $\times$ 1(Small Firm)	0.0123*** (0.00246)	0.0798* (0.0445)
1(Post) $\times$ 1(Treat) $\times$ 1(Small Firm)	-0.0415*** (0.0105)	-0.144* (0.0819)
Firm FE	✓	✓
Date FE	✓	✓
Observations	146104	146104
R-squared	0.118	0.624
Number of Clusters	13408	13408

Table IA.XV. Difference-in-differences: heterogeneity by presence of GitHub account

This table presents estimates of the effect of the FTX collapse on firm outcomes after the FTX shock. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post FTX* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. Standard errors in parentheses are clustered by firm. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,p} = \beta_0 + \beta_1(\text{post}_t \times \text{treat}_i) + \chi_i + \gamma_j \cdot 1(\text{firm}_j) + \delta_t \cdot 1(\text{start date}_t) + \varepsilon_{i,p}$$

	(1) %Δ Employees	(2) % External Share
1(Post) × 1(Treat)	-0.154*** (0.0476)	-0.562*** (0.138)
1(Post) × 1(Has GitHub)	0.0385** (0.0152)	-0.189** (0.0943)
1(Post) × 1(Treat) × 1(Has GitHub)	0.0935* (0.0479)	0.669*** (0.142)
Firm FE	✓	✓
Date FE	✓	✓
Observations	122174	122174
R-squared	0.110	0.640
Number of Clusters	11718	11718

Table IA.XVI. Difference-in-differences: firm outcomes

This table presents estimates of the effect of the FTX collapse on firm outcomes after the FTX shock. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post FTX* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. Standard errors in parentheses are clustered by firm. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,p} = \beta_0 + \beta_1(\text{post}_t \times \text{treat}_i) + \chi_i + \gamma_j \cdot 1(\text{firm}_j) + \delta_t \cdot 1(\text{start date}_t) + \varepsilon_{i,p}$$

	(1) Bankruptcy	(2) IPO	(3) M&A	(4) Secondary	(5) VC
1(Post) $\times$ 1(Treat)	0.0226 (0.0642)	0.00392 (0.00379)	0.0417 (0.0848)	0.0143 (0.148)	-2.146*** (0.443)
1(Has GitHub)	-0.0384** (0.0191)	0.0268*** (0.00519)	0.0405 (0.0306)	0.0592*** (0.0214)	-0.0527 (0.112)
1(Post) $\times$ 1(Has GitHub)	-0.0329 (0.0339)	-0.0143** (0.00571)	-0.0127 (0.0434)	0.0104 (0.0334)	0.405*** (0.125)
1(Treat) $\times$ 1(Has GitHub)	0.0576*** (0.0203)	0.0734*** (0.0163)	0.0127 (0.0614)	0.167 (0.121)	0.186 (0.361)
1(Post) $\times$ 1(Treat) $\times$ 1(Has GitHub)	-0.0405 (0.0661)	-0.0761*** (0.0172)	0.0156 (0.0902)	-0.00633 (0.157)	-0.753 (0.462)
Mean of DV	0.0773	0.0246	0.229	0.231	2.357
Date FE	✓	✓	✓	✓	✓
Observations	605466	605466	605466	605466	605466
R-squared	0.000280	0.000385	0.000168	0.00108	0.00490
Number of Clusters	11448	11448	11448	11448	11448

Table IA.XVII. Difference-in-differences: heterogeneity with continuous variables

This table presents estimates of the effect of the FTX collapse on employment growth and share of work complete by external developers. Treatment firms are those that received venture capital funding from investors who had previously invested in FTX. *Post FTX* is an indicator variable encoded as 1 for dates after June 2022 and 0 otherwise. The dependent variables are: (1) *Employee Growth* - the percentage change in employment; and (2) *External Share Volume* - the share of pull requests (by volume, or sum of code additions and deletions) contributed by external contributors. All specifications include firm and date fixed effects, with standard errors clustered at the firm and date level. Indicators for firm characteristics are formed as of December 31, 2021. This table is the continuous version of Table IA.XIV. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{Treat})_i \times 1(\text{bin}_i)) + \chi_i + \delta_t + \varepsilon_{it}$$

	(1) %Δ Employees	(2) % External Share
1(Post) × 1(Treat)	-0.126*** (0.0191)	0.103 (0.0765)
1(Post) × Age	0.000642*** (0.000114)	0.00121* (0.000636)
1(Post) × 1(Treat) × Age	0.00605*** (0.000925)	0.00368 (0.00441)
1(Post) × log(Headcount, 2021)	0.00125 (0.000942)	-0.0213* (0.0113)
1(Post) × 1(Treat) × log(Headcount, 2021)	0.00695** (0.00295)	-0.00200 (0.0191)
Firm FE	✓	✓
Date FE	✓	✓
Observations	91574	91574
R-squared	0.0935	0.604
Number of Clusters	5788	5788

Table IA.XVIII. Difference-in-differences: heterogeneity by VC

This table presents the same results from Table 4 but splits the treatment group by size of VC. 1(FTX VC Large) are firms that received funding from the five largest VCs that invested in FTX, based on number of deals. These are Sequoia Capital, Coinbase Ventures, Tiger Global Management, Digital Currency Group, and Lightspeed Partners. Column (1) and (3) are when the treatment group is only the largest 5 VCs while Column (2) and (4) are when the treatment group are all other treated VCs. All columns include firm and date fixed effects. Standard errors in parentheses are clustered by firm. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{Treat})_i) + \chi_i + \delta_t + \varepsilon_{it}$$

	%Δ Employees		Δ External Share	
	(1)	(2)	(3)	(4)
1(Post) × 1(FTX VC Large)	-0.0449*** (0.00539)		0.0888* (0.0510)	
1(Post) × 1(FTX VC Small)		-0.0459*** (0.00584)		0.119*** (0.0327)
Firm FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	111332	111332	111332	111332
Adj R-squared	0.0968	0.0974	0.637	0.638
Number of Clusters	10505	10505	10505	10505

Table IA.XIX. Difference-in-differences: heterogeneity by crypto-focused VC

This table presents the same results from Table 4 but splits the treatment group by size of VC. 1(FTX VC, Large Crypto Firms) are firms that received funding from the eight largest crypto-focused VCs that invested in FTX, based on number of deals. These are Coinbase Ventures, Digital Currency Group, Pantera Capital, Multicoin Capital, Paradigm, Hard Yaka, YZi Labs, Signum Capital, and HTX Ventures. Column (1) and (3) are when the treatment group is only the largest crypto VCs while Column (2) and (4) are when the treatment group are all other treated VCs. All columns include firm and date fixed effects. Standard errors in parentheses are clustered by firm. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{Treat})_i) + \chi_i + \delta_t + \varepsilon_{it}$$

	%Δ Employees		Δ External Share	
	(1)	(2)	(3)	(4)
1(Post) × 1(FTX VC, Large Crypto Funds)	-0.0572*** (0.00881)		0.198*** (0.0463)	
1(Post) × 1(FTX VC, Not Crypto Funds)		-0.0476*** (0.00551)		0.112*** (0.0324)
Firm FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	111332	111332	111332	111332
Adj R-squared	0.0969	0.0976	0.639	0.638
Number of Clusters	10505	10505	10505	10505

Table IA.XX. Difference-in-differences: placebo by other large VCs

This table presents the same results from Table 4 but assigns the treatment group to VCs that did not invest in FTX. In Column (1) and (3), 1(Placebo VC, Large Funds) are firms that received funding from Andreessen Horowitz, General Catalyst, Founders Fund, and Khosla Ventures, which are large VC funds based on fundraising amount. In Column (2) and (4), 1(Placebo VC, Many Deals) are firms that received funding from Y Combinator, Plug and Play Tech Center, Techstars, Alumni Ventures, and FJ Labs, which are the largest VC funds based on number of deals. All columns include firm and date fixed effects. Standard errors in parentheses are clustered by firm. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

$$y_{i,t} = \beta_0 + \beta_1(1(\text{Post})_t \times 1(\text{Treat})_i) + \chi_i + \delta_t + \varepsilon_{it}$$

	%Δ Employees		Δ External Share	
	(1)	(2)	(3)	(4)
1(Post) × 1(Placebo VC, Large Funds)	-0.0271*** (0.00565)		-0.0456 (0.0296)	
1(Post) × 1(Placebo VC, Many Deals)		-0.0242*** (0.00495)		0.0299 (0.0359)
Firm FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	111332	111332	111332	111332
Adj R-squared	0.0964	0.0965	0.636	0.636
Number of Clusters	10505	10505	10505	10505